UNIVERSITY OF LIVERPOOL

OXIDOREDUCTASES OF STROPHARIA AERUGINOSA

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

Maurycy Daroch

February 2011

ABSTRACT

Oxidoreductases of *Stropharia aeruginosa*.

Maurycy Daroch

Basidiomycete fungi are a rich source of oxidoreductases of potential biotechnological, industrial and environmental relevance. This thesis describes the continuation of work commenced by Dr J.K. Moore to identify a *Stropharia aeruginosa* haloperoxidase. An EST of a haloperoxidase-like protein of *S. aeruginosa* has been identified, along with important findings regarding another group of fungal enzymes, laccases. Two laccase isoenzymes were isolated and, apart from their spectroscopic properties, their characteristics are consistent with other laccases. Unusual spectroscopic properties of these proteins, characterised by the lack of a blue copper band at 600 nm, caused them to be annotated as yellow laccases. A novel approach using next generation sequencing technology and circular template PCR was used to isolate the cDNA sequences encoding these laccases. The deduced protein sequences of these proteins, the first sequences of yellow laccases, were used to create comparative models of their molecular structures on the basis of known crystal structures of laccases. Analyses of these models suggested that yellow laccases could posses more space around the T1 copper site than blue laccases, which could be suitable for binding of as yet unknown aromatic mediator molecules, which would expand their catalytic activity and explain their yellow colour.

CONTENTS

Chapter 1 – Introduction.

Chapter 2 – Materials and methods.

Chapter 3 – The pursuit of novel fungal haloperoxidase.

Chapter 4 – Isolation, purification and characterisation of *S. aeruginosa* oxidoreductases.

Chapter 5 – The pursuit of the genes of *S. aeruginosa* laccases using conventional methods.

Chapter 6 – The use of 454 sequencing in a pursuit of *S. aeruginosa* laccase genes.

Chapter 7 – The bioinformatic analysis of *S. aeruginosa* laccases.

Chapter 8 – Disscussion

References

Appendices

LIST OF FIGURES

Chapter 1 – Introduction.

Chapter 2 – Materials and methods.

Chapter 3 – The pursuit of novel fungal haloperoxidase.

Chapter 4 – Isolation, purification and characterisation of *S. aeruginosa* oxidoreductases.

Chapter 5 – The pursuit of the genes of *S. aeruginosa* laccases using conventional methods.

Chapter 6 – The use of 454 sequencing in a pursuit of *S. aeruginosa* laccase genes.

Chapter 7 – The bioinformatic analysis of *S. aeruginosa* laccases.

ACKNOWLEDGMENTS

First and foremost I would like to thank Andy Bates and Lesley Iwanejko for their guidance and support. Secondly, the thanks should go to Jean Wood without whom Lab G of Biosciences building would probably fall apart, and it would be quite difficult to do any experiments without her, keeping an eye on the lab and running it smoothly. Thirdly, a joint big thank you to all people that contributed in some parts of this project (in random order), Neil Hall (for allowing the use 454 sequencing system at marginal cost), Margaret Hughes (for doing the 454 libraries), Kevin Ashelford (for contig assembly), Mark Prescott (for running mass spec samples), Deborah Ward (for advice on 2D gels and proteomics), Mark Wilkinson (for advice and equipment for protein purification), Igor Morozov and Mark Caddick (for general help with molecular biology), Karolina Tkaczuk (for advice on comparative protein modelling), Leo Ciufo (for yeast work guidlines). Thirdly my acknowledgements go to project students: Mohammed M. Noor for hard work and creativity on *A. nidulans* AN 7823 putative haloperoxidase – it's a pity it did not work; Catherine A. Houghton for her very hard work on biochemical characterisation of *S. aeruginosa* laccases; Dan Forsey and Jay Hobson for their work on *P. fluorescens* esterase, I hope it wasn't a complete waste of time for them. Fourthly, the thanks go to my collaborators from the Technical University of Lodz, Marianna Turkiewicz and Tomasz Florczak. The special thanks should go to people from labs C, D, E and G (Magda, Kinga, Martyna, Kat, Ale, Yann, Paul F, Paul M, Paul E, Egon, Chris, Gaz, Pryank, Eli, Arnaud and many others) for socialising at the local and in more remote locations.

Last but not least I'd like to thank EU (FP 7 Marie-Curie Early Stage Training Fellowship) for research funding and a decent salary, Liverpool

Football Club and Rafa Benítez for entertainment in my free time and parents for my genetic pool and financial support few years ago. Thanks also go to my lovely fiancée 刘颖 for her patience and support in the writing up time. Significant acknowledgements should go to my former girlfriend Agata for her negative opinion about my scientific ambitions and skills; it's been a pleasure to prove her wrong. Many thanks go to vice-chancellor of the Technical University of Lodz prof. Stanislaw Bielecki, who encouraged me to do the PhD in the first place around five years ago.

No too much thanks goes to people that were in the lab but were not supposed to, self-touted lab DJs, once that don't fill the booking sheets and take out your samples from the thermocycler during the run, the once that nick your incubator whilst you're on holidays, and certain Tom and George that are not welcome in Liverpool, and the small blue bitter part of Merseyside.

ABBREVIATIONS

AaP = Agrocybe aegerita peroxidase

ABTS = 2,2′-azino-bis(3-ethylbenzthiazoline-6-sulphonic acid)

ATP = Adenosine-5′-triphosphate

dATP = Deoxyadenosine-5′-triphosphate

rATP = ATP

BLAST = Basic Local Alignment Search Tool

bp = base pairs

BSA = Bovine Serum Albumin

C-C = Carbon – Carbon bond

CBM = Cellulose Binding Motif

CBS = Centraalbureau voor Schimmelcultures

CHAPS = 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate

CPO = Chloroperoxidase

Da = Dalton

DEPC = Diethylpyrocarbonate

DMP = Dimethoxyphenol

DMSO = Dimethyl sulfoxide

DNA = Deoxyribonucleic acid

cDNA = complementary DNA

DTT = Dithiothreitol

E.C. = Enzyme Commission number

EDTA = Ethylenediaminetetraacetic acid

ELISA = Enzyme-Linked Immunosorbent Assay

EtBr = Ethidium bromide

FTP = File Transfer Protocol

g = rcf

HBT = 1-hydroxybenzotriazole

HIC = Hydrophobic Interaction Chromatography

HMM = Hidden Markov Model

HNC = High Nitrogen Content medium

IC_{50} = half maximal inhibitory concentration

IPG = Immobilized pH gradient

IPTG = Isopropyl β-D-1-thiogalactopyranoside

IUPAC = International Union of Pure and Applied Chemistry

kbp = kilo base pairs

LA = Luria-Bertani agar medium

LB = Luria-Bertani broth

LiP = Lignin Peroxidase

LMS = Laccase-Mediator System

MCD = Monochlorodimedone

MCS = Multiple Cloning Site

MM = Minimal Medium

M&M = Materials and Methods

MOPS = 3-(N-morpholino)propanesulfonic acid

MnP = Manganese Peroxidase

MSA = Multiple Sequence Alignment

NCBI = National Center for Biotechnology Information

NHE = Normal Hydrogen Electrode

NMR = Nuclear Magnetic Resonance

NN = Neutral Networks

OD = Optical density

PAGE = Polyacrylamide Gel Electrophoresis

pfu = plaque forming units

PCR = Polymerase Chain Reaction

PDB = Protein Databank

rcf = relative centrifugal force

RNA = Ribonucleic Acid

mRNA = Messenger RNA

tRNA = Total RNA

rpm = revolutions per minute

SA = Specific Activity

SD = Synthetic Drop-out medium

SD-U = Synthetic Drop-out medium without uracil

SDS = Sodium Dodecyl Sulphate

SSC = Saline Sodium Citrate buffer

T_m = Melting temperature

TE = Tris HCl-EDTA buffer

TMB = 3,3′,5,5′-Tetramethylbenzidine

U = Units

UV = Ultraviolet light

V = Volt

VIS = Visible Spectrum of light

X-gal = 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside

YPAD = Yeast extract Peptone Adenine sulphate Dextrose medium

Chapter 1

INTRODUCTION

1. Fundamentals of biocatalysis and biodegradation.

Biocatalysis and biodegradation are two related terms used to describe the use of biological catalysts, namely enzymes in a variety of applications. Enzymes can be used either as purified proteins or as whole cell catalysts. The fate of the end product determines to which of those two categories a given process will be classified. Biodegradation is associated with catabolism and the ability of an organism to decompose through the action of its enzyme(s) a substance into molecules that can enter central metabolic pathways of the organism, through an enzymatic cascade or a single enzymatic reaction. Biocatalysis, on the other hand, leads to product accumulation and is associated with both anabolism and catabolism (Wackett and Hershberger, 2001).

The classic example of biocatalysis is ethanol fermentation, a technology well known and extensively applied by mankind all over the world, dating back to the Neolithic Age, some 8000 years ago (Patrick, 1952). A thorough understanding of this technology, and the role of yeast in the process, was provided by Louis Pasteur in 1854 (Harden, 2009). The most obvious example of biodegradation is the carbon cycle of lignocellulose, in which an array of enzymes, of bacterial or fungal origin, largely glycoside hydrolases and oxidoreductases, break this complex polymer into simple compounds that can be utilized by the degrading organisms.

The two processes are inevitably linked with one another by the fact that the same enzyme can be used as a tool for both processes depending on the substrates and conditions used; for example nitrile hydratase of *Pseudomonas putida* can be used in the biodegradation of the herbicide bromoxynil (Vokounova et al., 1992) as well as in the synthesis of acrylamide from acrylonitrile (Rozzell, 1999). Both processes rely on the ability of the enzyme to hydrate the nitrile present in both starting compounds to an amide (Figure 1.1).

Both biocatalysis and biodegradation have attracted significant attention in recent years, due to a general drive to develop a more sustainable, ecologically friendly economy. Biodegradation is closely associated with bioremediation i.e. the use of the enzymatic machinery of microorganisms for the breakdown of toxic compounds like hyrocarbons, phenols and halogenated compounds, which can be found in excess in zones polluted by conventional industry (Mogensen et al., 2003). Biocatalysis offers the prospect of "green chemistry" (also sustainable chemistry); that is the substitution of conventional methods of chemical synthesis, traditionally involving the use of high energy input, toxic chemicals and hazardous solvents, with more eco-friendly, enzymatic catalysis, carried out at ambient temperatures with aqueous buffers. Enzymes also possess a unique characteristic compared with traditional catalysis, in that they are regio-, chemo- and stereospecific catalysts. These qualities can also be tuned by a variety of methods (Schulze and Wubbolts, 1999) that are unavailable for traditional catalysts such as protein engineering (rational and random protein mutagenesis and the prospect of *de novo* protein design (Arnold, 2001)). Enzymes therefore have a lot to offer for most synthetic reactions an organic chemist might require (Rozzell, 1999, Schoemaker et al., 2003) with few notable exceptions e.g. carbon cross coupling reactions catalysed by palladium.

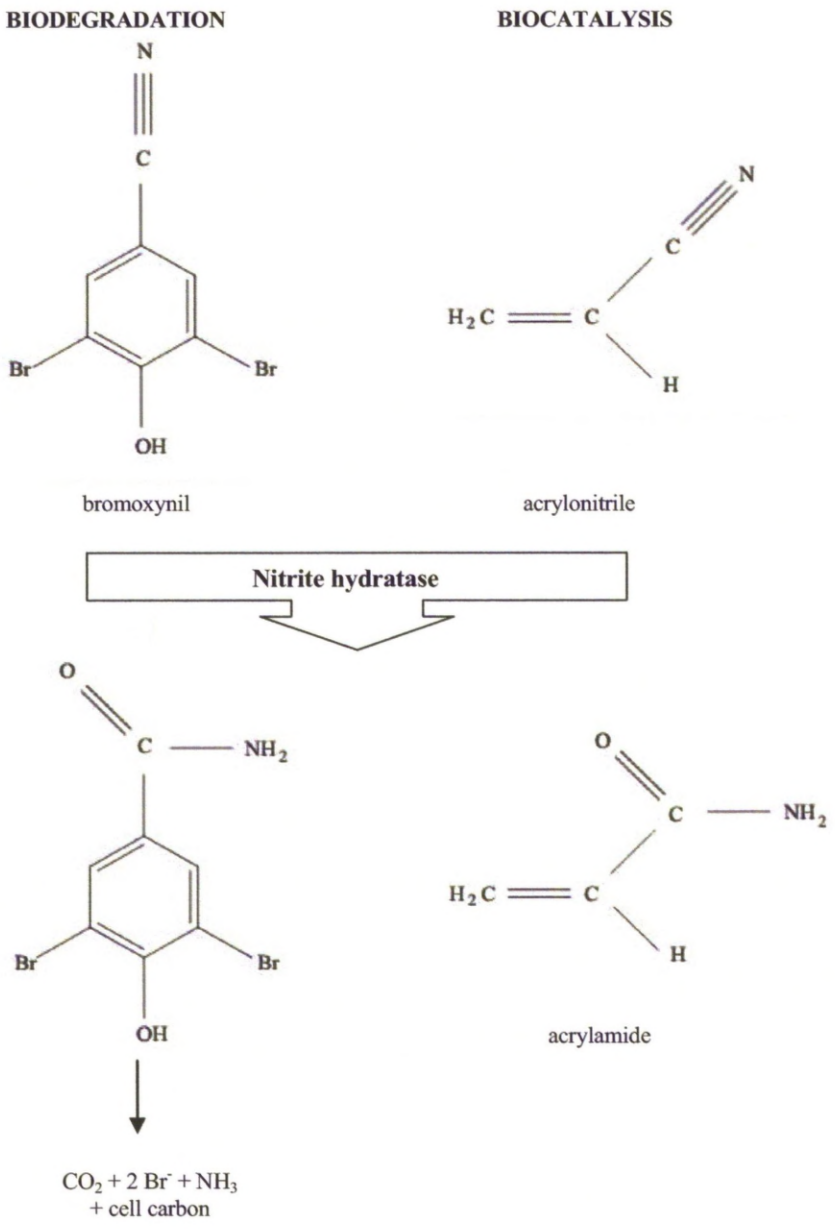


Figure 1.1 Biocatalysis and biodegradation of nitriles to amides by nitrile hydratase, adapted from Wackett and Hershbrg (2001).

The ability of nitrile hydratase to hydrate nitrile to amide is a reaction common to both biocatalysis and biodegradation. The biocatalytic route can be applied to produce acrylamide, whilst the biodegradative route decomposes bromonixyl to basic compounds such as ammonia and CO_2.

1.1. Biocatalysis and its applications.

Enzymes have been used to perform catalytic tasks on an industrial scale for a long time in the food industry, notably alcohol fermentation and cheese making. However, only at the end of nineteenth century could it be finally described as technology. In 1874 the first commercial enzymatic preparation of rennet, composed of natural chymosin and pepsin, was introduced to the market by Christian Hansen (Straathof and Adlercreutz, 2000). Now, almost 140 years on, enzymatic catalysis has flourished into a multi-billion Euros industry exploiting a whole range of nature's enzymatic potential. The proteins currently employed by biocatalysis nowadays ranges from "simple" hydrolases, mainly lipases, amylases and proteases to more sophisticated proteins like peroxidases, epoxidases and nitrile hydratases. Furthermore new avenues such as catalysis in supercritical fluids or ionic liquids make this technology even more promising for the future. Table 1.1 provides only a short summary of selected industrial applications of enzymes more information can be found in numerous books (Hou and Shaw, 2009, Bommarius and Riebel, 2004, Whittall and Sutton, 2009, Patel, 2006) and review papers focusing on different aspects of biocatalysis(Kirk et al., 2002, Zaks, 2001, Zhao et al., 2002, Zhao and Van Der Donk, 2003, Schoemaker et al., 2003).

Table 1.1. Selected industrial applications of enzymes, adapted from Kirk, Borchert et al. (2002)

Industry	**Enzyme class**	**Application**
Detergent	Protease	Protein stain removal
	Lipase	Lipid stain removal
	Cellulase	Cleaning, colour clarification, anti-redeposition (cotton)
Starch and fuel	Amylase	Starch liquefaction and saccharification
	Glucose isomerase	Glucose to fructose conversion
	Xylanase	Viscosity reduction
Food	Protease	Milk clotting, infant formulas (low allergenic), flavour
	Lactase	Lactose removal (milk)
	Pectinase	Fruit-based products
Baking	Amylase	Bread softness and volume, flour adjustment
	Xylanase	Dough conditioning
	Lipoxygenase	Dough strengthening, bread whitening
Animal feed	Phytase	Phytate digestibility – phosphorus release
	β - Glucanase	Digestibility
Beverage	Pectinase	De-pectinization, mashing
	Laccase	Clarification (juice), flavor (beer), cork stopper treatment
	Amylase	Juice treatment, low calorie beer
Textile	Cellulase	Denim finishing, cotton softening

	Pectate lyase	Scouring
	Catalase	Bleach termination
	Peroxidase	Excess dye removal
	Laccase	Bleaching
Pulp and paper	Lipase	Pitch control, contaminant control
	Protease	Biofilm removal
	Xylanase	Bleach boosting
	Cellulase	De-inking, drainage improvement, fiber modification
Fats and oils	Lipase	Transesterification
	Phospholipase	De-gumming, lyso-lecithin production
Organic synthesis	Lipase	Resolution of chiral alcohols and amides
	Acylase	Synthesis of semisynthetic penicillin
	Nitrilase	Synthesis of enantiopure carboxylic acids
Leather	Protease	Unhearing, bating
	Lipase	De-pickling
Personal care	Amyloglucosidase	Antimicrobial (combined with glucose oxidase)
	Glucose oxidase	Bleaching, antimicrobial
	Peroxidase	Antimicrobial

1.2. Biodegradation and its applications.

Large scale industry produces a variety of structurally different xenobiotic contaminants resistant to biodegradation by strains traditionally found in soil. These substances include: polycyclic aromatic hydrocarbons, halogenated phenols, polychlorinated biphenyls, trichloroethylene, trinitrotoluene, phthalate esters, petroleum products, azo dyes and surfactants. Due to the complex nature of xenobiotics, biodegradation of these compounds to basic substances requires whole metabolic pathways rather than an action of individual enzymes. Decomposition with specialized bacterial and fungal strains provides the best solution from both the ecological and economic point of view (Diaz, 2007) and is an absolute requirement if we are about to develop a sustainable economy with low environmental impact. More information about this wide area of enzymatic catalysis can be found in books (Erkurt, 2010, Diaz, 2007, Singh and Ward, 2004) and review papers (Peng et al., 2008, Pandey et al., 2007, Nyahongo et al., 2005, Sinha et al., 2009, Mogensen et al., 2003).

2. Lignocelullose and its degradation.

Lignocellulose, a key building block of all plants, is the biggest source of renewable energy on Earth, storing carbon and solar energy for millions of years in a form of living plants and fossil fuels deposited underground. This complex biopolymer is composed of three major components: cellulose, lignin and hemicellulose and very small amounts of other components like pectins, proteins, ash or lipids (Dashtban et al., 2009). The main individual components of lignocellulose, cellulose and lignin are the most abundant natural polymers in the ecosystems (Lundell et al., 2010).

Degradation of all of these components requires the synergistic actions of many classes of enzyme specific for individual components of the biopolymer. Among the best degraders of lignocellulose are the fungi, mainly basidiomycetes, and the anaerobic bacteria (Dashtban et al., 2009, Perez et al., 2002, Beguin and Lemaire, 1996).

2.1. Lignocelullose structure and function.

Unlike other natural polymers like starch, glycogen, proteins and nucleic acids lignocellulose is a structure of incredible resistance to degradation, often described as biomass recalcitrance. The main structural component of biomass is cellulose – a highly crystalline polymer composed of D-glucose, just like easily hydrolysable starch and glycogen. The key difference between these two structures is the glycosidic bond, which is oriented in a β (1→4) fashion opposite to starch that is mainly composed of an α (1→4) linked D-glucose with an occasional α (1→6) branchings. The opposite orientation of this bond causes glucose units to form a dense network of hydrogen bonds, strengthened by Van der Waals forces and hydrophobic interactions that make it very difficult to break down.

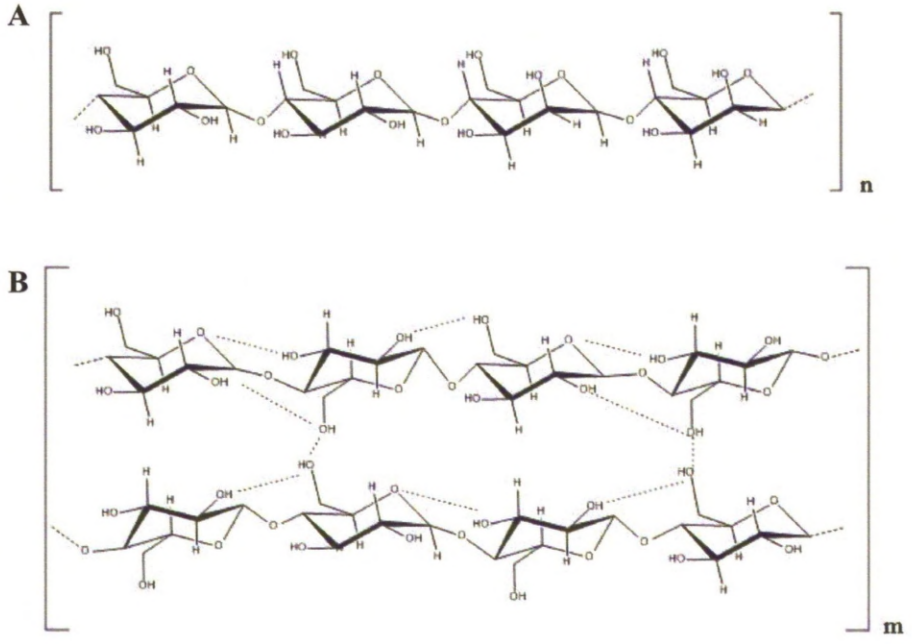


Figure 1.2 Chemical structures of starch and cellulose (Roberts, 1986).

The amylose fraction of starch (A) is composed of α (1→4) linked *D*-glucose units. The average length (n) of an amylose chain is 300-3000 glucose units, every 24-30 glucose units an α (1→6) branching occurs – this structure is called amylopectin.
The cellulose structure (B) is composed of β (1→4) linked *D*-glucose units without any branching. The length of individual chain (cellobiose) can range from 300 to as high as 10,000 glucose units (m). This structure is highly stabilized by internal hydrogen bonds (indicated with dotted lines).

Another complex carbohydrate present in lignocellulose is hemicellulose, this carbohydrate is much more complex in composition, being formed by a variety of pentoses, hexoses and sugar derived acids. The main components of this structure are: *D*-xylose, *D*-mannose, *D*-galactose, *D*-glucose, *L*-arabinose, 4-*O*-methyl-glucuronic acid, *D*-galacturonic and *D*-glucuronic acids. Sugars are connected with each other *via* β (1→4) and β (1→3) glycosidic bonds. The structure of hemicellulose is branched and is relatively easily hydrolysable, providing that appropriate enzymatic machinery, able to accept such a variety of substrates, is available. Lignin is the third main component of lignocellulose. It is

synthesised *via* the reaction of free radicals and is composed of three main components: dimethoxylated, monomethoxylated and non-methoxylated phenyl propanoid units that are derived from corresponding p-hydroxycinnamyl alcohols (Figure 1.3) (Martinez et al., 2005, Baucher et al., 1998). These components form syringyl, guaiacyl and p-hydroxyphenyl components of lignin that are connected *via* C-C, aryl ether linkages predominantly being aryl-glycerol and β aryl structures (Sanchez, 2009). Lignin is linked to the sugar components of lignocellulose forming a physical seal protecting cellulose and hemicellulose from microbial attack and oxidative stress. Lignin also has a major function to play as a structural support for the plant (Leonowicz et al., 1999, Sanchez, 2009).

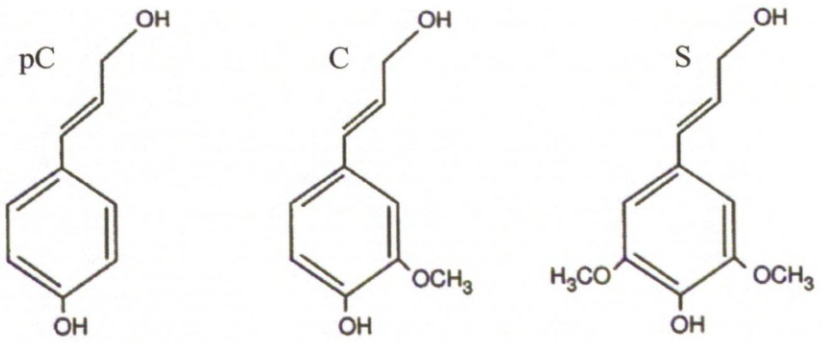


Figure 1.3 Lignin precursors. Adapted from Baucher (1998).

Major lignin building blocks i.e. *p*-hydroxyphenyl, guaiacyl and syringyl units are composed from corresponding *p*-hydroxycinnamyl alcohols: *p*-coumaryl (pC), corifenyl (C) and syringyl (S).

The compositions, varies considerably from one plant species to another, or even between different organs. The lignin content can vary from 0 % for some leaves and cotton cobs to over 30 % for softwood stems and bamboo, and around 40 % for nut shells (Sanchez, 2009). The contribution of hemicellulose varies from around 25 % for rice and grasses to 80 % for some leaves. The amount of cellulose usually falls between 20 and 50 % for most species and organs (Sanchez, 2009). It should be also noted that various species and organs contain different types of lignin (ratio of individual types of phenylpropanoid units and linkages), various composition of hemicelullose (different sugar molecules present), and distinct degrees of cellulose crystallinity (Yoshida et al., 2008, Le Ngoc Huyen et al., 2010).

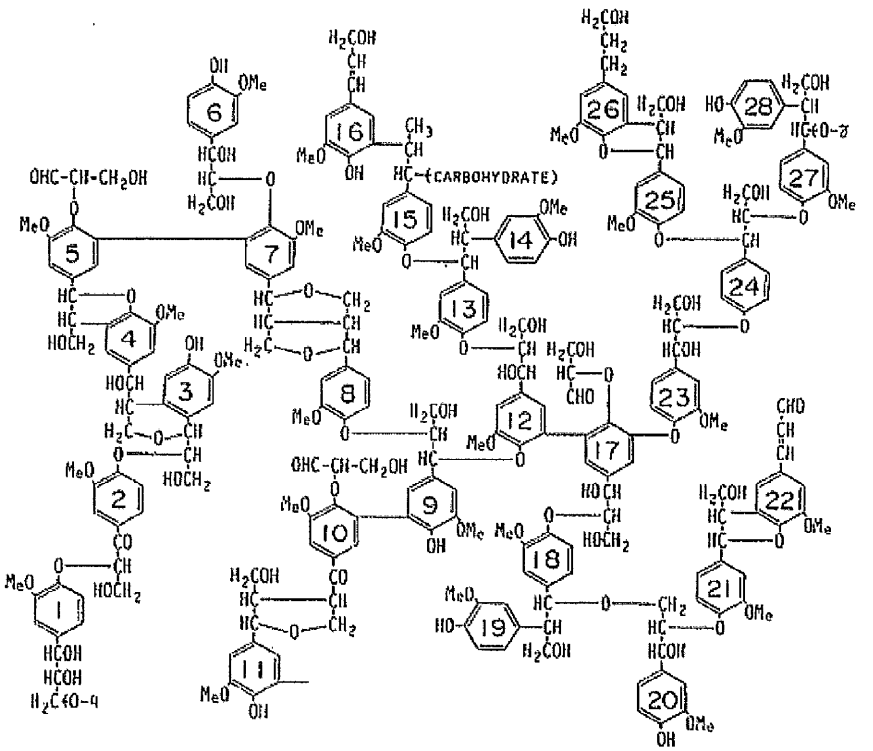


Figure 1.4 Structure of softwood lignin. Adapted from Sakakibara (1977). The structure of softwood lignin containing all lignin *p*-hydroxyphenyl, guaiacyl and syringyl units and typical linkages.

2.2. Enzymatic degradation of lignocellulose.

To degrade lignocellulose Nature has developed two distinct systems associated with the type of respiration of microorganisms. Aerobic organisms, fungi and some bacteria rely on secretion of individual enzymes, acting in synergy, to decompose lignocellulose, whereas anaerobic bacteria and few anaerobic fungi found in the gastrointestinal tract of herbivores have evolved a complex extracellular system called the cellulosome, which is composed of scaffolding proteins and many bound enzymes (Doi and Kosugi, 2004). Recently published work shows, however that *Limnoria quadripunctata*, a wood-borer is capable of lignocellulose degradation without the assistance of symbiotic bacteria (King et al., 2010).

2.2.1. Aerobic degradation of lignocellulose.

Fungi, especially basidiomycetes are the best lignocellulose degraders on Earth. Four main types of fungal degradation of lignocellulose can be distinguished: white rot (further divided into simultaneous rot and selective delignification), brown rot, soft rot and stain fungi. White and brown rot is generally performed by basidiomycetes, wheras the latter two types mainly by ascomycetes and deuteromycetes (Martinez et al., 2005). Fungal degradation of lignocellulose requires the simultaneous action of an array of enzymes specific for the individual components of the heterogeneous substrate.

To deal with the cellulose component of lignocellulose, fungi secrete an array of cellulases acting on different regions of carbohydrate fibres. The degradation of cellulose is started by the action of endocellulases (endo-1,4-β-glucanases) that hydrolyse internal β (1→4) glycosidic bonds of amorphous cellulose "chopping" them into smaller fragments and creating a new terminal of cellobiose. Those fragments can than be

processed by two classes of cellobiohydrolases (exo-1,4-β-glucanases) acting from either the reducing (hemi-acetal C1 carbon of glucose) (class I) or non-reducing (hydroxyl C4 carbon of glucose) (class II) end of the carbohydrate chain. The final hydrolysis into glucose units requires the action of β-glucosidase (Perez et al., 2002, Sanchez, 2009).

The main component of hemicellulose is xylan that requires two main enzymes for its hydrolysis; an endo-1,4-β-xylanase and xylan 1,4-β-xylosidase. The former generates oligosaccarides that are subsequently hydrolysed by the latter into xylose molecules. The heterogeneity of hemicellulose requires a whole array of accessory enzymes for non-xylose molecules and, among others, includes: xylan esterases, ferulic and *p*-coumaric acid esterases, arabinofuranosidases and glucuronosidases (Kirk and Cullen, 1998).

Lignin is the most recalcitrant polymer to decompose and is degraded by fungi *via* oxidative reactions, sometimes described as enzymatic combustion. The term characterises a chaotic nature of this process. The first step involves an electron transfer from the enzyme to lignin. After this initial oxidation the degradation is performed by the action of unstable lignin degradation intermediates generated by the first step. In the end the breakdown of the polymer occurs *via* fission of carbon-carbon or carbon-oxygen bonds of the side chains or through cleavage of the aromatic rings takes place (Marzullo et al., 1995). Three main enzymatic components are thought to be responsible for the degradation of lignin: two types of peroxidises and laccases (Figure 1.5). The main enzyme responsible for lignin cleavage is lignin peroxidase (LiP), a high redox potential enzyme capable of oxidizing substrates up to +1.4 V (Kersten et al., 1990). LiP is capable of oxidizing both phenolic and non-phenolic components of lignin (linkages). Another enzyme is manganese peroxidase (MnP), which has a lower redox potential than LiP. The MnP mechanism of action relies on the oxidation of Mn^{2+} to Mn^{3+} which is stabilised by a strong organic chelator, such as oxalate or malonate (Youngs et al., 2000), and is then released into

the environment. MnP cannot oxidise non-phenolic units of lignin. Both MnP and LiP must be coupled with a generator of free H_2O_2 that is required as a co-substrate for the catalysis. The H_2O_2 is usually provided by an oxidase e.g. glyoxal oxidase or aryl alcohol oxidase. The last enzyme traditionally associated with lignin biodegradation is laccase. This protein belongs to the family of multicopper oxidases and, unlike peroxidases, does not require any cofactors for its action, its redox potential does not exceed +0.7 V and it is only specific for the phenolic regions of lignin. All of the oxidoreductases are usually present as multiple isoforms, allelic variants, splice variants or glycoforms each, showing slightly different substrate preference (Perez et al., 2002, Kirby, 2005, Martinez et al., 2005).

To summarise, fungal degradation of lignocellulose requires the simultaneous action of many enzymes, peroxidises coupled with oxidases and laccases for lignin breakdown, and another large array of cellulases and hemicellulases for carbohydrate hydrolysis. In nature it is a time consuming process that sustains the natural carbon cycle and is performed by a relatively narrow group of organisms.

2.2.2. Anaerobic degradation of lignocellulose.

The main characteristic that distinguishes the anaerobic system from its aerobic counterpart is the presence of a scaffold that coordinates spatially the action of enzymes, mainly cellulases and hemicellulases. The scaffold is built from pairs of cohesins and dockerins that connect all the individual components into a functional entity. The overall structure and function of the carbohydrate hydrolases remains the same with respect to known proteins from aerobic systems. The major difference is that instead of the cellulose binding motifs (CBM), found in free enzymes, a cohesion-dockerin structure is found. The celulosome is directed towards the substrate by one scaffoldin-associated CBM (Bayer et al., 2004). Currently no lignin-degrading enzymes have been found in the bacterial cellulosome.

The only additional enzymes are hemicellulases, chitinases and lichenases that allow the organism to retrieve more cellulose for degradation (Bayer, 2000).

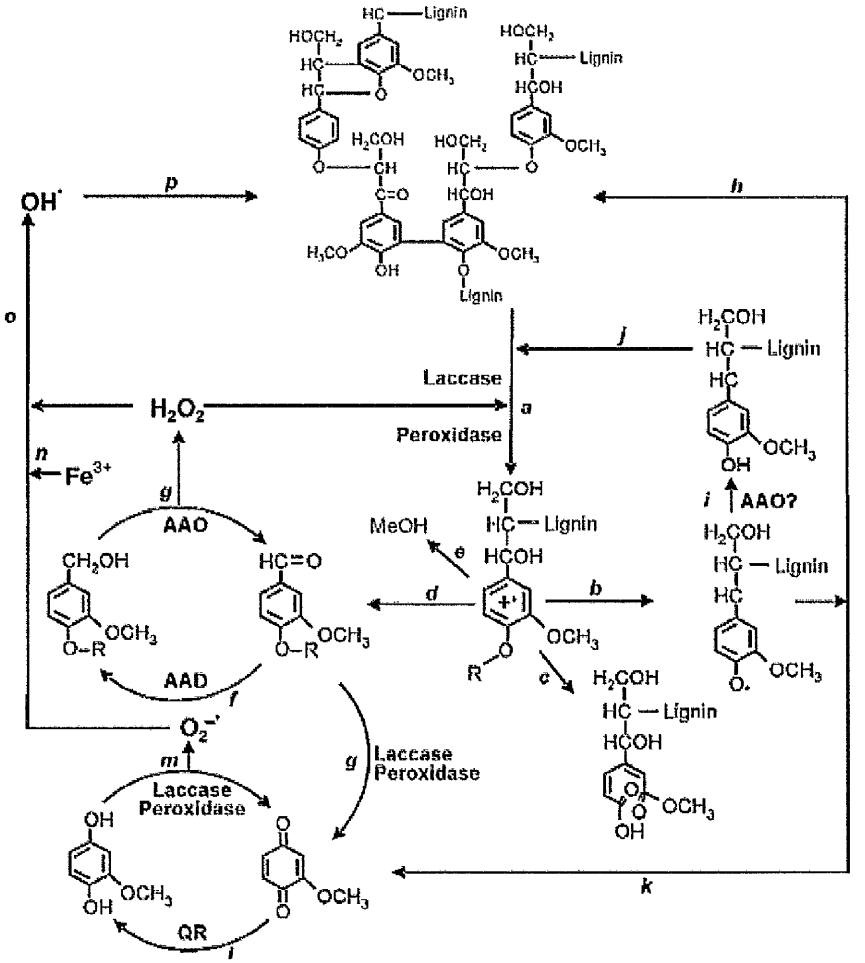


Figure 1.5 Simplified scheme of fungal degradation of lignin. Adapted from Martinez (2005).

Aromatic radicals (a) are generated by LiP, MnP and laccase through C4 ether breakdown (b), aromatic ring cleavage (c), Cα-Cβ breakdown (d) and demethoxylation (e). H_2O_2 required by peroxidases is generated through aryl alcohol oxidase-aryl alcohol dehydrogenase couple. The products of C4 ether breakdown (b) can however repolymerise to lignin (h) if they are not reduced by oxidases to phenolic compounds (i) or become reoxidised (j). These radicals can be also subjected to Cα-Cβ breakdown to quinones (k). Quinones can contribute to oxygen activation in redox cycling (l, m) resulting in reduction of ferric ion in wood (n) by superoxide cation radical or semiquinone radicals yielding hydroxyl radical $OH^{\cdot}$ (o) that can attack lignin (p) in the initial stages of its decay.

3. Oxidoreductases and their applications.

The oxidoreductases (E.C. 1) are a major class of enzymes catalysing reactions of oxidation-reduction i.e. transfer of electrons from electron donor to electron acceptor. Potential industrial applications of oxidoreductases are vast, ranging from synthetic chemistry through biosensors and diagnostics to biomass degradation and environmental biotechnology (May, 1999).

3.1. Peroxidases.

Peroxidases (E.C. 1.11.1.7) are widely distributed in nature and have been identified among plants, animals and microorganisms. Peroxidases are mainly haem enzymes that contain iron (III) protoporphyrin IX (ferriprotoporphyrin IX), as the prosthetic group and are divided into two major groups: plant and mammalian (Regalado, 2004). Plant peroxidases contain plant ascorbate peroxidases and other classical plant peroxidases, fungal peroxidases and yeast cytochrome C peroxidase. Mammalian peroxidases include: myeloperoxidase, thyroid peroxidase, lactoperoxidase and prostaglandin H synthetase (Welinder, 1992).

Peroxidases are enzymes that catalyse the oxidation of an organic molecule at the expense of peroxide that acts as an electron acceptor in a two-electron oxidising reaction. The general scheme of the reaction is presented in Figure 1.6 (Martinez, 2002, Rydberg et al., 2004, Battistuzzi et al., 2010).

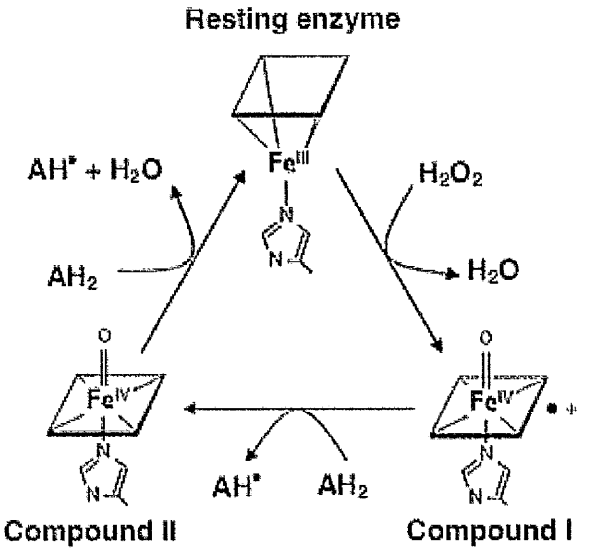


Figure 1.6 The catalytic cycle of peroxidases. Adapted from Battistuzzi (2010).

The catalytic cycle starts with the binding of H_2O_2 to the Fe^{3+} of the porphyrin which becomes deprotonated to form the hydroperoxide intermediate. The next step involves reprotonation on the distal oxygen and dissociation of water to create the highly reactive compound I. Compound I then reacts with the substrate by abstracting its electron and forming another intermediate, compound II. This complex then abstracts an electron from another molecule of the substrate, as a result the catalytic iron returns to the Fe^{3+} state after taking two protons and dissociating a water molecule.

Because of their high oxidative capabilities peroxidases have been long considered enzymes of very high industrial potential. The first commercial applications of peroxidases were found in colorimetric tests such as antibody labelling for ELISA or diagnostic tests for detection of glucose or uric acid and other compounds. More recently fungal peroxidases have been considered as an environmentally friendly method of industrial dye decolourization, polymer synthesis, wood pulp bleaching and delignification (Regalado, 2004). The latter application seems to have attracted considerable attention in recent years due to a general drive towards more environmentally friendly processes and the development of biofuels and biorefining (Shia et al., 2009).

Despite their numerous advantages, peroxidases seem to suffer a number of drawbacks that seriously limit their biotechnological potential. First and foremost these enzymes require exogenous hydrogen peroxide as a co-substrate. They also suffer considerable inactivation through the oxidation of the porphyrin ring by H_2O_2. So far, no real solution has been found to prevent this process, although it can be limited by the step-wise addition of hydrogen peroxide to the reaction mixture. Some advances have been made through directed evolution, however only the prolongation of peroxidase half-life was achieved (Morawsky et al., 2001).

The other major drawbacks limiting the application of peroxidases are both the difficult production and improvement of these proteins. To be cost efficient these proteins should be produced in heterologous systems that would also allow relatively straightforward protein engineering. Unfortunately peroxidases seem to be notoriously difficult to express in reasonable quantities due to a combination of factors: the host organism must be capable of ensuring proper folding, correct coordination of the prosthetic group, inclusion of structural calcium ions and proper glycosylation. It is very difficult to fulfil all of these criteria and achieve high expression levels at low cost. So far the best way to produce peroxidases heterologously is the use of filamentous fungi e.g. *Aspergillus*, however protein engineering is limited in this expression system (Ayala et al., 2008).

3.2. Haloperoxidases.

Haloperoxidases (E.C. 1.11.1.10) are enzymes that catalyse the incorporation of a halogen atom into an organic molecule in a two-step process at the expense of hydrogen peroxide. There are three different types of haloperoxidases: haem, vanadium and non-metal haloperoxidases (Littlechild, 1999). The best studied enzyme thus far is chloroperoxidase of *Caldariomyces fumago* (Morris and Hager, 1966) a haem-containing

enzyme. Its natural function is chlorination in the biosynthesis of an antimicrobial compound caldariomycin. The catalytic mechanism of chloroperoxidases is a two step process; firstly a molecule of hypohalous acid is formed from oxidation of a halide by hydrogen peroxide (enzymatic step). The acid then attacks an organic compound in a nonenzymatic reaction through electrophillic substitution (van Pee et al., 2006). The process is presented on Figure 1.7:

Enzymatic reaction:

$H_2O_2 + X^- + H^+ \rightarrow HOX + H_2O$

Nonenzymatic reaction:

$HOX + AH \rightarrow AX + H_2O$

X = Cl, Br or I

Figure 1.7 The mechanism of action of haem haloperoxidases.

The first step of the catalysis is an oxidation of the halide by hydrogen peroxide to create a molecule of hypohalous acid. The next step involves the attack of the hypohalous acid on a molecule of substrate through electrophilic substitution.

In vitro, haloperoxidases show similar activities to cytochrome P450, catalyzing reactions of epoxidation, sulphoxidation and various oxidations. What differentiates haloperoxidases from P450s is the lack of a cofactor requirement for NAD(P)H (Ullrich and Hofrichter, 2005).

Quite recently, two interesting findings were published: the discovery of a haloperoxidase in brown-rot basidiomycete *Agrocybe aegerita* (Ullrich and Hofrichter, 2005) and evidence of lignin chlorination by a litter decomposing fungus *Curvularia inaequalis* (Ortiz-Bermudez et al., 2007), those two facts can suggest that biological halogenation may be an ubiquitous process that can somehow be associated with lignocellulose biodegradation and the carbon cycle.

3.3. Laccases.

3.3.1. Introduction to laccases.

Laccases (EC 1.10.3.2) are oxidative enzymes belonging to the family of multicopper oxidases. They catalyse the reaction of one electron oxidation of a reducing substrate coupled with a four electron reduction of molecular oxygen to water, using four copper atoms according to the following scheme:

$$2\,H^{+} + 4\,S + O_2 \rightarrow 4\,S^{+\cdot} + 2\,H_2O$$

Figure 1.8 The mechanism of action of laccases.

Laccases are enzymes that catalyse sequential one-electron oxidations of four substrate molecules coupled with a four electron reduction of molecular oxygen to water.

Laccases are one of the oldest-known enzymes and were discovered as a component of resin ducts of a tree *Rhus venicifera* by Yoshida at the end of 19th century (Yoshida, 1883). They are widely found in fungi, plants and have been also identified in bacteria and insects (Riva, 2006). Their ability to perform catalysis using molecular oxygen yielding water as a by product make them the most "green" and "eco-friendly" oxidreductases around.

These green oxidoreductases are usually monomeric globular proteins of molecular weight around 60-70 kDa and isoelectric point of 4.0, the extent of glycosylation usually ranges from 10 to 25 %. They contain a four nuclear copper cluster in three distinct sites in the catalytic sites called T1 (one atom), T2 (one atom), T3 (two atoms) (Figure 1.9). T1 is the catalytic, so-called "blue copper" (Giardina et al., 2010, Baldrian, 2006). Their natural function has been traditionally associated with lignin degradation, stress defence, morphogenesis and fungal plant-pathogen/host interactions.

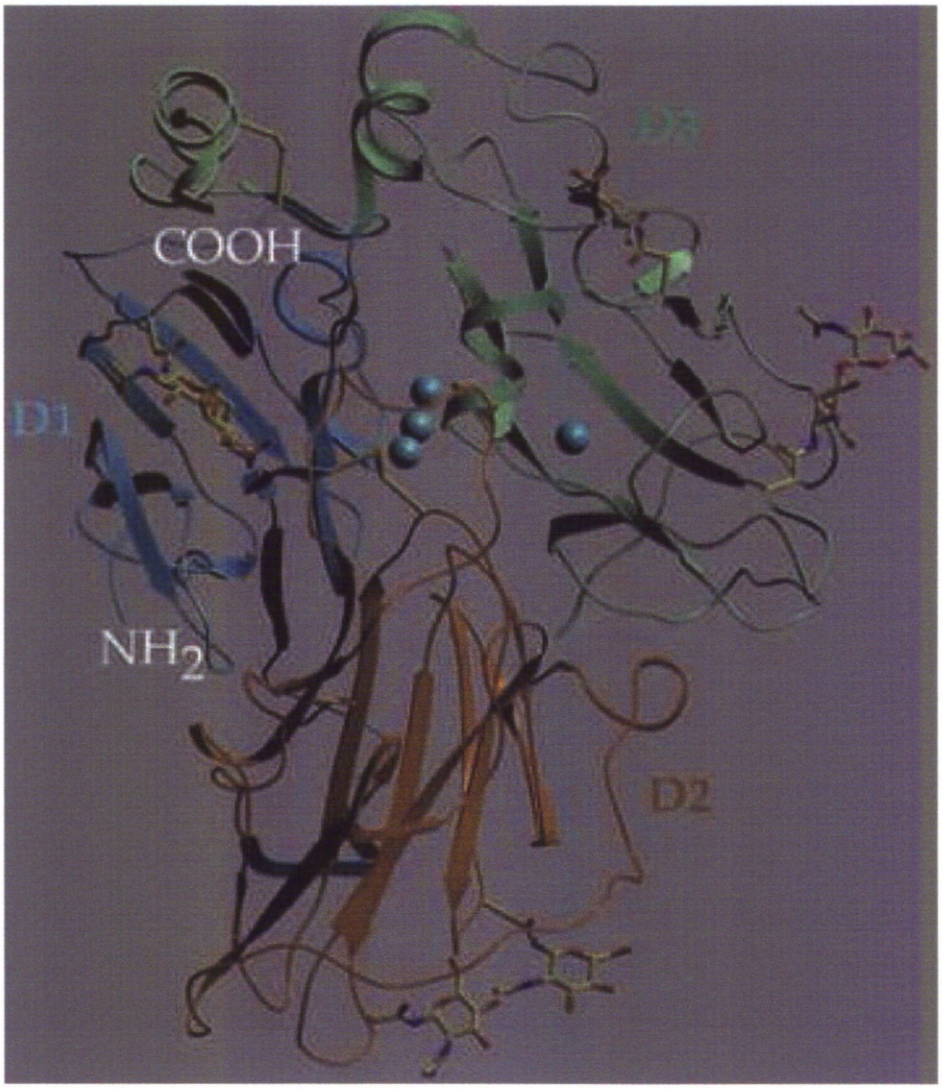


Figure 1.9 Crystal structure 1GYC showing *Trametes versicolor* laccase structure. Adapted from Piontek (2002)

Laccase of *T. versicolor* is a monomeric protein composed of three domains. The interface of D2 and D3 contains catalytic "blue" copper site T1, trinuclear copper site is at the interface between three domains. Laccase structure contains seven *N*-acetyl glucoseamine at five distinct *N*-glycosylation sites.

3.3.2. Blue and yellow laccases.

The vast majority of laccases described in the literature belong to the typical blue laccases. The characteristic blue colour of these proteins results from the T1 copper site, and charge transition from the sulphur atom of cysteine to Cu^{2+} atom. This charge transfer gives rise to a very characteristic absorption band of 612 nm. The first unusual laccase lacking this feature was isolated from the fungus *Panus tigrinus*. The biochemical characterisation showed that the enzyme not only exhibits similar substrate specificities as blue laccases but also contains all four copper atoms per molecule which negates some suggestions that copper depleted protein was isolated (Leontievsky et al., 1997a). Since then several yellow laccases have been isolated (Min et al., 2001, Leontievsky et al., 1997b, Pozdnyakova et al., 2004, Edens et al., 1999) and characterised with various levels of detail; however to date no gene or structural information is available that would reveal the difference between blue and yellow laccases. It is postulated by Leontievsky that yellow laccases are essentially identical to blue laccases, the only difference being the presence of a lignin-derived mediator in the catalytic site that causes blue colour to disappear and expands the substrate specificity of an enzyme, making it capable of also oxidizing non-phenolic lignins (Leontievsky, personal communication).

Several other findings have showed that laccases are a more diverse group of proteins than it was first expected, to date white laccases, in which some copper atoms were substituted by zinc and iron have been described (Palmieri et al., 1997). Other work suggests that it is possible to replace metals in the four sites of multicopper oxidases to tune their catalytic abilities (Nakamura and Go, 2005).

3.3.3. Biotechnological application of laccases.

The lack of cofactors and limitations associated with peroxidase-catalysed reactions make laccases a very interesting group of enzymes that could replace peroxidases in some industrial processes. The major drawback of laccases is their considerably lower redox potential compared to peroxidases. The maximal theoretical potential of a laccase resulting from thermodynamic threshold is +1.2 V vs. NHE (normal hydrogen electrode) compared to the potential of +1.9 V shown by peroxidases (Ayala et al., 2008). To tackle the problem two main strategies have been employed, directed evolution of proteins for higher redox potential (Zumarraga et al., 2008), and the use of mediators that expand the substrate specificity of a laccase (Wells et al., 2006). Both strategies can work at the same time leading to even greater improvement of laccase biotechnological potential. The mediators, low molecular weight compounds, can be of natural or synthetic origin and serve as an oxidation reaction carrier between the enzyme and the substrate (Riva, 2006). One of the interesting features of these molecules is that some of them can be oxidised twice before carrying the redox charge to the substrate, this makes the laccase-mediator system competitive with redox potential of peroxidases. The concept of a laccase mediator system is presented below.

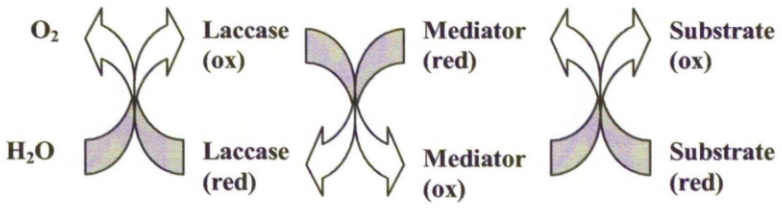


Figure 1.10 Concept of laccase-mediator system.

A laccase molecule oxidises a molecule of mediator (once or more) which then acts as an electron transfer mediator to the molecule of substrate. That allows laccases to expand their substrate specificity to less accessible substrates of higher redox potential.

Due to its natural function in lignocellulose degradation, laccases have been successfully applied to wood processing and in the processes of biopulping and biobleaching, to limit the usage of hazardous chlorine-based chemicals (Widsten and Kandelbauer, 2008). In the late 1990s the process Lignozyme® was introduced. It is an enzymatic system of biobleaching based on the mixture of peroxidases and a laccase mediator system (LMS) of laccase/HBT (1-hydroxybenzotriazole). Another field where laccases show remarkable potential is wastewater treatment and industrial/textile dye decolourisation. Laccases show very good results in degradation of various azo and anthraquinone dyes commonly employed by industries (Chagas and Durrant, 2001, Michniewicz et al., 2008, Pereira et al., 2007). Another field where laccases are used is the bioremediation of polycyclic aromatic hydrocarbons, which has been successfully achieved by many groups using different laccases (Alcalde et al., 2002, Pozdnyakova et al., 2006, Juhasz and Naidu, 2000). Some promise was also shown in the food industry where it was found that laccases' ability to cross link biopolymers make them suitable for the baking industry (Selinheimo et al., 2006). Laccases have long been considered interesting tools for synthetic organic chemistry, but application and development of LMS and the design of mediators has made this area even stronger (Kunamneni et al., 2008a). The laccase mediator system finds applications in the synthesis of aldehydes from alcohols and primary benzyl amines, thioaldehydes from thioalcohols and oximes and nitrones from hydroxylamines Last but not least, the aromatization reaction from dehydro aromatics has also been performed (Wells et al., 2006). Other applications involve polymerization reactions to produce polycatechol (Aktas and Tanyolas, 2003) and polymeric polyphenols; as well as oxidation of carbohydrates. Currently development is focused on increasing laccase tolerance for organic solvents through directed evolution (Kunamneni et al., 2008a, Alcalde et al., 2005) and development of designer mediators to expand the synthetic abilities of these proteins (Wells et al., 2006).

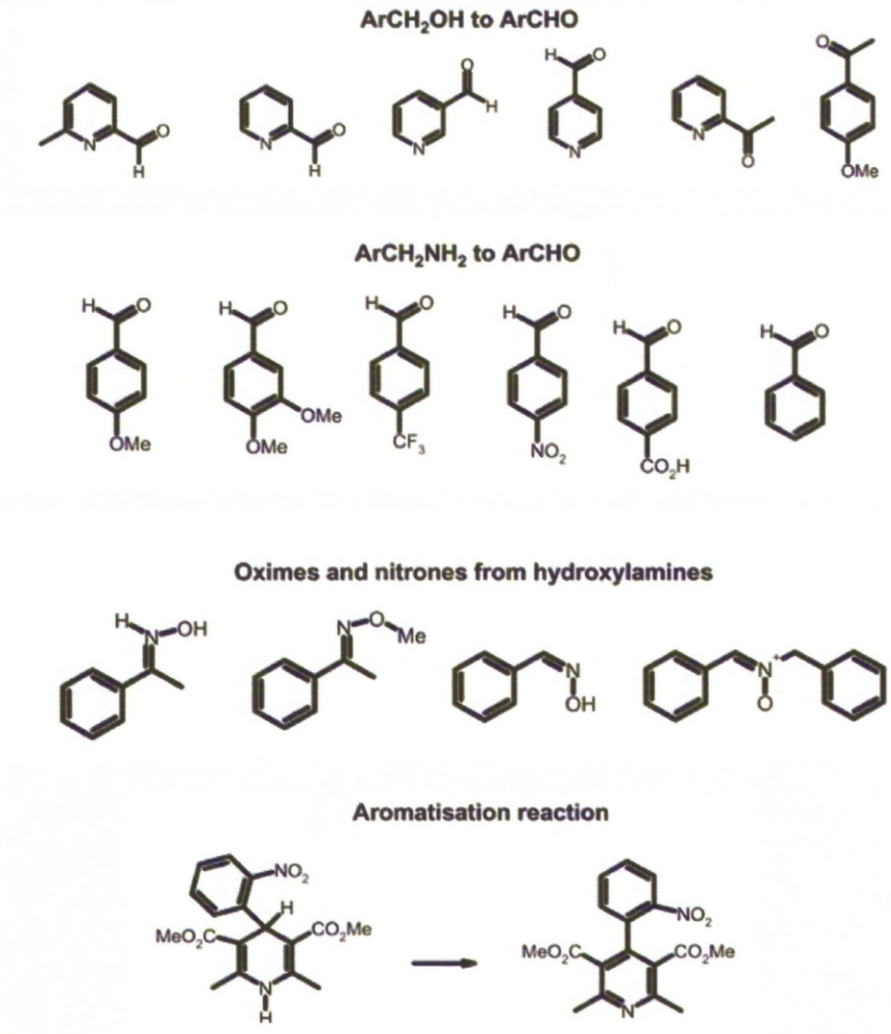


Figure 1.11 Typical synthetic reactions that can be obtained with laccase-mediator system. Adapted from Wells (2006)

Laccase-mediator system offers synthetic organic chemistry a number of useful applications including: oxidation of alcohols to aldehydes, oxidation of primary benzyl amines to imines and their hydrolysis to aldehydes, oxidation of hydroxylamines to oximes and nitrones and aromatisation of dehydro aromatics

Laccases are much better enzymes for heterologous production than peroxidases (Ayala et al., 2008). They were successfully expressed in filamentous fungi (Tellez-Jurado et al., 2006), yeast (Colao et al., 2006, Bulter et al., 2003) and even bacteria (Martins et al., 2002). The greatest yields were obtained for filamentous fungi, but the possibility of high throughput screening methods given by bacteria and yeast makes protein engineering much easier. A good review about large scale laccase production in bioreactors is available (Couto, 2007) highlighting the advantages of solid state fermentation for industrial scale production of laccases from filamentous fungi.

4. Aims and objectives.

The original aim of this work was to follow Jonathan K. Moore's PhD work titled "Pursuit of novel haloperoxidase for asymmetric biocatalysis" (Moore, 2007). Two approaches to find a novel haloperoxidase were tried. The first was focused on bioinformatic identification of a protein similar to *Caldariomyces fumago* chloroperoxidase. The AN7823 protein from *Aspergillus nidulans* was identified using bioinformatic approaches and attempts were made to express it heterologously in *Eshcerichia coli*. In parallel, attempts were made to isolate the *Stropharia aeruginosa* chloroperoxidase-like protein reported by Moore. Unfortunately, neither approach was successful. Attempts were then focused on several laccases of *S. aeruginosa* that were identified as members of an unusual group of yellow laccases. It was postulated that these proteins show interesting features regarding their spectroscopic properties and substrate specificity. Several similar proteins have been purified, their DNA sequences elucidated and provisional biochemical and modelling analyses have been performed.

Chapter 2

MATERIALS AND METHODS

1. Strains used.

1.1. Fungal strains.

Stropharia aeruginosa CBS 839.87 strain was purchased from Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands.

Geomyces sp. P7 was obtained from the pure culture collection of the Institute of Technical Biochemistry of Technical University of Lodz, Poland.

Aspergillus nidulans wild type strain WT *bi A1* was kindly provided by Magdalena Moś and Mark Caddick, University of Liverpool.

1.2. Bacterial strains.

Lambda ZAP work cells:

Escherichia coli XL1-Blue MRF′ [Stratagene]
Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1recA1 gyrA96 relA1 lac* [F′ *proAB lacI*qZΔ*M15* Tn*10* (Tet^r)]

Escherichia coli SOLR™ [Stratagene]

e14–(McrA–) Δ(*mcrCB-hsdSMR-mrr*)*171* *sbcC* *recB* *recJ* *uvrC* *umuC::*Tn*5* (Kanr) *lac* *gyrA96* *relA1* *thi-1* *endA1* λR [F′ *proAB* *lacI*qZΔ*M15*] Su– (nonsuppressing)

DNA cloning cells:

Escherichia coli XL1-Blue [Stratagene]

recA1 *endA1* *gyrA96* *thi-1* *hsdR17* *supE44* *relA1* *lac* [F′ *proAB* *lacI*qZΔ*M15* Tn*10* (Tet^r)]

Escherichia coli SoloPack® Gold [Stratagene]

Tet^r Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173* *endA1* *supE44* *thi-1* *recA1* *gyrA96* *relA1* *lac* Hte [F′ *proAB* *lacI*qZΔ*M15* Tn*10* (Tet^r) Amy Cam^r]

Escherichia coli TOP10F′ One Shot® [Invitrogen]

F′ {*lac*Iq Tn*10* (Tet^R)} *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ*ΔM15 Δ*lacX*74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G

Escherichia coli JM 109 [Promega]

*rec*A1, *end*A1, *gyr*A96, *thi*, *hsd*R17 (r_K^-,m_K^+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F′, *tra*D36, *pro*AB, *lac*IqZΔM15]

Escherichia coli DH5α [Bioline]

F- *deo*R *end*A1 *rec*A1 *rel*A1 *gyr*A96 *hsd*R17(r_k^-, m_k^+) *sup*E44 *thi*-1 *pho*A Δ(*lac*ZYA *arg*F)U169 Φ80*lacZ*ΔM15 λ^-

Escherichia coli SoloPack® for pSCB vector [Stratagene]

Whole genotype information not provided.

endA1 *recA1* Str^R

Protein expression cells:

Escherichia coli BL21 (DE3) [Novagen]
F$^-$ *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3)

Escherichia coli BL21 (DE3) pLysE [Novagen]
F$^-$ *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pLysE (Cam^R)

1.3. Yeast strains.

Saccharomyces cerevisiae INVSc1 [Invitrogen]
MATa, his3D1, leu2, trp1-289, ura3-52

Saccharomyces cerevisiae BJ5465
Ade2-1, ura3-52, trp3-11, pep-his, poblΔ1, 6R can1-100, GAL
The protease-deficient BJ5465 strain was kindly provided by Miguel Alcalde (CSIC Madrid, Spain)

2. Growth of organisms.

2.1. Growth media.

Stropharia aeruginosa was cultivated in High Nitrogen Content medium (HNC): (2 % (w/v) glucose [BDH], 0.5 % (w/v) mycological peptone [Oxoid], 0.2 % (w/v) yeast extract [Sigma], 0.1 % (w/v) KH_2PO_4 [BDH], 0.05 % (w/v) $MgSO_4$ [BDH], 6 ppm (w/v) NaCl [BDH]). Solid HNC medium used for routine replantation of fungus was supplemented with 1.5 % (w/v) of agar [Difco].

Aspergillus nidulans was grown in minimal medium containing 10 % (w/v) glucose [BDH], 0.05 % (w/v) KCl [Sigma], 0.05 % (w/v) $MgSO_4 \cdot 7 H_2O$ [BDH] , 0.15 % (w/v) KH_2PO_4 [BDH], 40 ppm (w/v) $Na_2B_4O_7 \cdot (10 H_2O)$ [BDH], 400 ppm (w/v) $CuSO_4$ [BDH], 800 ppm (w/v) $(FePO_4)_2 \cdot (5 H_2O)$ [BDH], 800 ppm $MnSO_4$ [BDH], 800 ppm (w/v) Na_2MoO_4 [Sigma], 0.8 % (w/v) $ZnSO_4$ [BDH], pH 6.5.

2.2. Growth conditions.

2.2.1. Fungal strains.

Agar cultures of *Stropharia aeruginosa* were routinely replanted every 2 months onto fresh HNC agar medium. The 1 cm diameter agar plugs from these plates were used to inoculate liquid cultures and for subculturing. Liquid cultures were cultivated stationary at room temperature. After four weeks; fungal biomass was harvested and used for nucleic acid isolation and culture medium used for protein isolation by ammonium sulphate precipitation.

Aspergillus nidulans was cultivated in suspension culture over 3 days at 37 °C, 150 rpm.

2.2.2. Bacterial strains.

For cloning and protein expression purposes all *E.coli* strains were cultivated overnight at 37 °C, 200 rpm in LB (1% (w/v) tryptone [Difco] 1% NaCl [BDH] 0.5 % (w/v) yeast extract [Sigma] pH 7.0) or LA (LB medium supplemented with 1.5 % (w/v) agar [Difco]) medium supplemented with an appropriate antibiotic if necessary up to a concentration of 100 μg mL^{-1}. If IPTG and X-gal were required for colour screening they were added at concentrations recommended by the strain manufacturer. Protein expression was induced by addition of IPTG to a final concentration of 50 μM (unless stated otherwise) during exponential growth phase.

For lambda phage work, strains were cultivated in LB medium supplemented with 10 mM $MgSO_4$ and 0.2 M (w/v) maltose and plated on NYZ agar plates (0.5 % (w/v) NaCl [BDH], 0.2 % (w/v) $MgSO_4$ 7 H_2O [BDH], 0.5 % (w/v) yeast extract [Sigma], 1 % NZ amine [Sigma], 1.5 % (w/v) agar [Difco], pH 7.5) as recommended by the manufacturer.

2.2.3. Yeast strains.

Saccharomyces cerevisiae strains were grown in a non-selective YPAD medium: 1 % (w/v) bacto yeast extract [Difco], 2 % (w/v) bacto peptone [Difco], 2 % (w/v) glucose [BDH], 0.008 % (w/v) adenine hemisulphate [Sigma].

After transformation with a URA3 gene containing plasmid, yeast were grown on Synthetic Dropout medium without uracil (SD-U) containing: 0.67 % (w/v) Yeast Nitrogen Base without amino acids [Sigma], 0.192 % (w/v) yeast amino acid mix without uracil [Sigma], 2%

glucose [BDH] or galactose [Sigma] (for protein expression). Agar media were supplemented with 1.5 % (w/v) agar [Difco]

3. DNA and RNA manipulations.

3.1. Genomic DNA isolation.

Mycelia from 4 week stationary growth were ground in liquid nitrogen to a fine powder. Powder (approx. 0.5 g) was extracted with 1 mL DNA extraction buffer (300 mM Tris-HCl, 375 mM NaCl, 37.5 mM EDTA, 3% (v/v) Sarkosyl pH 8.5) and 1 mL of Tris-HCl-saturated phenol pH 8.0 [Sigma]. After gentle shaking, 1 mL of chloroform:isoamyl alcohol (24:1) mixture was added and slurry was gently mixed for at least 10 min. The mixture was transferred to microcentrifuge tubes and centrifuged at 13,000 g for 15 min. The aqueous (upper) layer was collected and treated with 10 µL RNaseA [Fermentas] (10 mg mL^{-1}) for at least 1 h at 37 °C. After incubation the solution was extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1) [Sigma] and once with chloroform:isoamyl alcohol (24:1). The aqueous (upper) phase was removed after each step *via* centrifugation (13,000 g, 15 min.). DNA was then precipitated overnight with 0.11 volume of 3 M NaOAc and 2 volumes of ethanol. After incubation DNA was pelleted by centrifugation, washed twice with 70 % (v/v) ethanol and resuspended in 30 µL of TE (10 mM Tris-HCl, 1 mM EDTA pH 7.5).

3.2. Total RNA isolation.

For RNA work all solutions were treated with 0.1 % DEPC-diethylpyrocarbonate [Sigma] to inactivate RNAses. Mycelia were frozen in liquid nitrogen and ground into a fine powder. The ground cells (approx. 0.5 g) were then added to tubes containing 1 mL of lysis buffer and 1 mL

of Tris-HCl-saturated phenol, pH 5.0. The tubes were centrifuged for 10 min at 13,000 g, the upper phase removed and placed into a clean tube with an equal amount of Tris-HCl-saturated phenol pH 5.0. Tubes were centrifuged as above. The upper phase was removed and placed into a 1.5 mL tube with an equal volume of 5 M LiCl. The tubes were then incubated overnight at -20 °C. Following incubation, the tubes were centrifuged for 20 min. at 13,000 g. The subsequent pellet was washed with 180 µL of 70 % (v/v) ethanol, followed by centrifugation for 1 min at 15,000 g. The resulting pellet was then dissolved in 100 µL H_2O then precipitated overnight at -20 °C with 0.1 volume of 3 M NaOAc and 2 volumes of ethanol. After incubation, tubes were centrifuged for 30 min at 13,000 rpm. The pellet was washed in 180 µl of 70% (v/v) ethanol and centrifuged for 1 min at 13,000 g. The pellet was then dissolved in an appropriate volume of TE and stored in -80 °C.

3.3. mRNA isolation.

Messenger RNA was isolated from total RNA pool using Dynabeads® mRNA DIRECT™ Kit [Invitrogen] as stated in the manufacturer's protocol.

3.4. Nucleic acid quantification.

The concentration of DNA and RNA was measured with Nanodrop [Thermo Scientific] using 1 µL of a sample per measurement.

3.5. Nucleic acid concentration by evaporation.

If necessary the DNA and RNA samples were concentrated using a vacuum concentrator [Flowgen].

3.6. Gel extraction of DNA.

Gel extraction of DNA was performed with QIAEX II gel extraction kit [Qiagen] according to the manufacturer's instructions.

3.7. Agarose gel electrophoresis of DNA.

Agarose gel electrophoresis of DNA was performed in Fisherbrand horizontal electrophoresis gel tanks [Fisher Scientific] using agarose at a concentration of between 0.7 to 2.0 % (w/v) in 1 x TAE (40 mM Tris-Acetate; 1 mM EDTA) buffer pH 8.0. Gels were run at 50 to 200 V depending on the size of the apparatus used until the bromophenol blue dye reached ¾ of the gel length. 1 in 5 volumes of loading buffer was added to DNA samples prior to loading into the agarose gel. DNA was stained by the addition of ethidium bromide to the molten agarose at a final concentration of 0.5 μg/mL and visualised under UV. Hyperladder I or Hyperladder V [Bioline] was used as molecular weight markers (typical image of ladder is shown in Appendix 2).

3.8. Agarose gel electrophoresis of RNA.

Agarose gel electrophoresis of RNA was performed in Fisherbrand horizontal electrophoresis gel tanks [Fisher Scientific] using agarose at a concentration of 1.3 % (w/v) in 3.7 % (v/v) formaldehyde, 1 x MOPS (20 mM MOPS; 5 mM sodium acetate; 1 mM EDTA) buffer pH 8.0. Gels were run at 100 V until the bromophenol dye reached ¾ of the gel length. The running buffer was 1 x MOPS buffer pH 7.0. The RNA samples were mixed with the freshly made sample buffer containing: 0.12 μg EtBr; 300 μL of 10 x MOPS pH 8.0; 80 μL 37% formaldehyde and 900 μL formamide at the ratio 3 to 2 and heated at 65 °C for 150 s before snap cooling on ice. After the run RNA was visualised under UV.

3.9. Polyacrylamide gel electrophoresis of DNA.

Polyacrylamide gel electrophoresis of small DNA fragments was performed in a Mini Protean 3 System [Bio-Rad] using a 5 % acrylamide gel (29:1 acrylamide:bisacrylamide ratio) in 1 x TBE (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA) pH 8.3. The running buffer was 1 X TBE and the samples were run until the bromophenol dye reached the end of the gel. After the run the gel was stained with 0.5 $\mu g \cdot mL^{-1}$ EtBr in H_2O and visualised under UV. Hyperladder V [Bioline] was used as a molecular weight marker (typical image of ladder is shown in Appendix 2).

3.10. DNA sequencing.

All DNA samples were sequenced by Genomes Enterprise Ltd., John Innes Centre, Norwich, UK.

3.11. High-throughput 454 DNA pyrosequencing.

The hi-throughput sequencing of *S. aeruginosa* and *Geomyces sp. P7* cDNAs prepared with SMART™ kit was performed by Dr Margaret Hughes from the Centre for Genomic Research of the University of Liverpool using a 454 GS-FLX Titanium pyrosequencer [Roche]. The assembly of contigs from raw data was performed by Dr Kevin Ashelford from the Centre for Genomic Research.

3.12. Lambda ZAP cDNA library construction.

The cDNA library of *S. aeruginosa* was constructed using a ZAP-cDNA Synthesis Kit [Stratagene] from 5 μg of pure mRNA according to the manufacturer's instructions with two minor modifications:

- The mRNA secondary structures were disrupted by heating the mixture to 95°C and subsequent cooling on ice before the first strand synthesis.
- The methylmercury hydroxide relaxation step was omitted.

3.13. Non standard techniques used with *S. aeruginosa* λZAP library.

3.13.1. Whole library amplification.

Around 10^5 pfu were plated on 15 cm Petri dishes of NYZ medium in a form of liquid culture in NZY Top agar according to the manufacturer's instructions. In total approximately 10^6 pfu were used in the mass excision protocol as stated in Stratagene manual (p. 41). After excision *E. coli* SOLR cells were transfected with excised phagemids and plated onto LA medium supplemented with ampicillin. Around 10^6 colonies were washed off from plates using LB medium supplemented with ampicillin (100 μg mL^{-1}) and incubated with vigorous shaking (250 rpm) for 2 hours at 37 °C. After this preincubation an equal fresh volume of LB medium supplemented with ampicillin (100 μg mL^{-1}) was added and the culture was incubated for an additional hour at 37 °C, 200 rpm and used for protein production through induction with 50 μM IPTG (final concentration) or isolation of plasmid DNA with QIAfilter Plasmid Midi Kit [Qiagen]

3.13.2. Library transfer to *S. cerevisiae.*

The directional cDNA of *S. aeruginosa* constructed with the λZAP library kit was ligated to an EcoRI/XhoI-digested and dephosphorylated pYES-MCS vector. The ligated library (100 ng) was transformed into 6 vials of SoloPack® Gold supercompetent cells [Stratagene] according to the manufacturer's instructions plated on LA-amp plates, amplified essentially as stated in 3.12.1. and used for transformation of *S. cerevisiae* with a total of 5 μg of plasmid DNA.

3.14. SMART™ cDNA synthesis.

The SMART™ cDNA of *S. aeruginosa* was synthesised using a SMART™ PCR cDNA Synthesis Kit [Clontech] from 1 μg of pure mRNA according to the manufacturer's instructions.

3.15. Plasmid DNA isolation.

Small and large-scale plasmid preps were performed using QIAprep Spin Miniprep Kit [Qiagen] and QIAfilter Plasmid Midi Kit [Qiagen] respectively according to the manufacturer's instructions.

3.16. Restriction digest.

Restriction digests of DNA were performed with restriction enzymes [Fermentas] according to the manufacturer's instructions.

3.17. Dephosphorylation of plasmid vector.

If required, after restriction digest vectors were dephosphorylated with FastAP thermosensitive alkaline phosphatase [Fermentas] according to the manufacturer's instructions.

3.18. Ligation of DNA fragments.

DNA fragments were ligated with T4 DNA ligase [Promega] according to the manufacturer's instructions.

3.19. Phosphorylation of DNA fragments.

DNA fragments were phosphorylated with T4 polynucleotide kinase [Promega] according to the manufacturer's instructions.

3.20. Recircularisation of DNA fragments.

DNA fragments of interest were prosphorylated and subsequently ligated in the same reaction with T4 polynucleotide kinase [Promega] and T4 DNA ligase [Promega] according to following protocol:

2 μL of polynucleotide kinase buffer [Promega]

2 μL of 10 mM ATP [Stratagene]

2 μL of polynucleotide kinase (10 U)

X μL DNA (100-200 ng)

Y μL sterile distilled water to 20 μL

The sample was incubated at 37 °C for 1h and than subsequently heat inactivated at 65 °C for 20 min. After heat-inactivation the following components were added:

2 μL of polynucleotide kinase buffer [Promega]

2 μL of 10 mM rATP [Stratagene]

2 μL of T4 DNA ligase (6 U)

14 μL sterile distilled water

Following reaction was vortexed and incubated at 4 °C overnight.

3.21. Site-directed mutagenesis.

Site-directed mutagenesis was performed using a QuikChange Site-Directed Mutagenesis Kit [Stratagene] according to the manufacturer's instructions.

3.22. A tailing of PCR product.

When necessary for TA cloning the PCR products were A-tailed with GoTaq polymerase [Promega], a typical reaction mixture of 10 μL contained:

2 μL of 5 x reaction buffer

2 μL 1 mM dATP [Roche]

1 μL GoTaq polymerase

5 μL of gel purified blunt ended PCR product

3.23. Synthetic oligonucleotides.

Custom designed oligonucleotides were ordered form Invitrogen and are summarised in Table 2.1.

Table 2.1. Synthetic oligonucleotides used for *S. aeruginosa* project

Primer name	Primer sequence 5′ → 3′	Length [bp]	GC [%]	T_m [°C]	T_{ea}* [°C]	Chapter
GPNAN_FWD	GTTCTTTAATAGTGGACTC	19	36.8	44.7	45	V 4.1.
GPNAN_REV	TAAICGTAANCCNG	14	42.9	43.3 (40.0–47.7)	45	V 4.1.
VWNVIV_FWD	CGTTCTTTAATAGTGGACTC	20	40.0	47.8	45	V 4.1.
VWNVIV_REV	TGCTAITGTAAGGTITG	17	35.3	47.5	45	V 4.1.
VVNATAV_FWD	GTNGTIAATGCNACIGC	17	47.1	53.5 (50.1–57.4)	50	V 4.1.
VVNATAV_REV	TTTCCTGCGTTATCCCCTG	19	52.6	54.8	50	V 4.1.
InsAMP_FWD	ATTTCCATTCGCCATTCAGG	20	45.0	53.0	50	V 4.1.
InsAMP_REV	AGGCACCCCAGGCTTTACAC	20	60.0	60.0	50	V 4.1.
PolyT_REV	CGAGTTTTTTTTTTTTTTT	22	13.6	43.9	45	V 4.1.
Mauro502IvP_Fwd	CTGTCAATAACTTTACGTTTGTGCCCC	27	44.4	58.3	62	VI 6.
Mauro502IvP_Rev	AGCCAAGCTGATGTTGAAGGC	22	50	58.5	62	VI 6.
Mauro518IvP_Fwd	CCGAAAATCCTCCCGCTTAACGC	23	56.5	60.6	62	VI 6.
Mauro518IvP_Rev	CCAAGGACCAGAGTTGTCCGTTTC	24	54.2	59.7	62	VI 6.

Table 2.1. Synthetic oligonucleotides used for *S. aeruginosa* project (continued).

Primer name	Primer sequence 5′ → 3′	Length [bp]	GC [%]	T_m [°C]	T_{ea}* [°C]	Chapter
Mauro641IvP_Fwd	GCGGTCATTGCACCACTC	18	61.1	57.0	62	VI 6.
Mauro641IvP_Rev	ACGTCAACACGATTGCTACG	20	50	55.5	62	VI 6.
Mauro968IvP_Fwd	TCATATCGACTGGCACCTTG	20	50	54.5	62	VI 6.
Mauro968IvP_Rev	GGTGTGCCCATGAAGATGG	19	57.9	56.3	62	VI 6.
Mauro1153IvP_Fwd	GCCGTTAATCAAGGTTGCGTCTGC	24	54.2	60.9	62	VI 6.
Mauro1153IvP_Rev	CGTCTTGTTTCCCTCTCCTGCG	22	59.1	60.0	62	VI 6.
Mauro1577IvP_Fwd	AGACTGATAATGCCGGTCCTTGG	23	52.2	58.8	62	VI 6.
Mauro1577IvP_Rev	AGACTGATAATGCCGGTCCTTGG	22	54.5	58.5	62	VI 6.
Mauro3441IvP_Fwd	GCAAAGGACGTTATGCAGAGGACC	24	54.2	60.0	62	VI 6.
Mauro3441IvP_Rev	GACGATAAGTGGTCCCGGAAATGAC	25	52.0	59.2	62	VI 6.
Mauro4378IvP_Fwd	GTTGTCTACGACCGCCATGATCC	23	56.5	59.7	62	VI 6.
Mauro4378IvP_Rev	CCTTGAAGCCAGTGACAAGTGGG	23	56.5	60.2	62	VI 6.
MauroHalo1192IvP_F	CCTTGTACCCAGTCAAGCAAAGCC	24	54.2	60.3	62	VI 6.
MauroHalo1192IvP_R	CGAGCATCGACTATCGCCTTCC	22	59.1	59.6	62	VI 6.

Table 2.1. Synthetic oligonucleotides used for *S. aeruginosa* project (continued)

Primer name	Primer sequence 5′ → 3′	Length [bp]	GC [%]	T_m [°C]	T_{ea}* [°C]	Chapter
LacMid1_FWD	CYATGYTGMYGAGYACWAGYATTCATTGGC	30	53.6	63.0 (56.5-69.4)	60	VI 7.
LacMid1_REV	CCAWGGMCCMRTTRTCNGTYTYRAASCG	28	46.7	61.0 (55.2-66.8)	60	VI 7.
LacNterm1_FWD	ATGCAAGTGCAAGCTCTACTCCGATTC	27	48.1	60.8	60	VI 7.
LacNterm2_FWD	ATGCTTTCCCTCTCAGCCATTGTTGG	26	50.0	61.4	60	VI 7.
LacCterm1_REV	TCAATCAAAAGTTTGTGGAGGTAAGGC	27	40.7	57.7	60	VI 7.
LacCterm2_REV	TTAGTCCCTGCAGAGATTCTCCCAAGC	27	51.9	61.7	60	VI 7.
LacCterm3_REV	TCAGGTGAATACCTGAGGAGGCAGAG	26	53.8	61.2	60	VI 7.
LacCterm4_REV	TCATAAATGAGGGCACAGATTATCCCAAGC	30	43.3	60.2	60	VI 7.

GC –GC content,
T_m – calculated melting temperature; calculation according to Oligo Analyzer http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/
T_{ea} – experimental annealing temperature,
Chapter – where to find in the Thesis

Table 2.2. Synthetic oligonucleotides used for *A. nidulans* AN 7823 project, and cloning adaptor for pYES 2.1 vector.

Primer name	Primer sequence 5′ → 3′	Length [bp]	GC [%]	T_m [°C]	T_{an}* [°C]	Chapter
Cloningadaptor_F	GGAATTCCCGGGGATCCATGGATATCGGTACC TCGAGCA	39	56.4	69	45	V 2.4.1.
Cloningadaptor_R	GCTCGAGGTACCGATATCCATGGATCCCCGGG AATTCCA	39	56.4	69	45	V 2.4.1.
AsperClon_F	CAATGTTGCTCAAGAGCATCCA	22	45.5	55.9	56	III 4.
AsperCloneNoSP_F	ATGGAGCTGGACTTTGAACAGTGGCA	26	50	62.4	56	III 4.
AsperCloneNoSP_NcoI_F	AGCCATGGAGCTGGACTTTGAACAGTGGCA	30	53.3	66.5	56	III 4.
AsperCloneNoSP_KpnI_R	ATTGGTACCTCAAGCACGAAACG	23	47.8	57.7	56	III 4.
MutaNcoI_1_F	CGATCAAGTACTTCCGAACGATGGTCGACCCG CGC	35	60	68.7	65	III 5.1.
MutaNcoI_1_R	GCGCGGGTCGACCATCGTTCGGAAGTACTTGA TCG	35	60	68.7	65	III 5.1.
MutaNcoI_2_F	CCTTTGACCCGCTGCCGTGGCAGCGGCCTCCC ATCTGG	38	71.1	75.5	65	III 5.1.
MutaNcoI_2_R	CCAGATGGGAGGCCGCTGCCACGGCAGCGGG TCAAAGG	38	71.1	75.5	65	III 5.1.

3.24. Polymerase chain reaction DNA amplification.

All PCR amplifications were performed in heated lid thermocycler DNA Engine DYAD Peltier [Bio-Rad] using following polymerases:

3.24.1. Taq polymerase RedTaq Ready Mix.

Amplification with RedTaq [Sigma] was performed essentially according to manufacturer's conditions, a typical reaction mixture of 50 µL contained:

0.5 µL of forward primer (100 µM)

0.5 µL of reverse primer (100 µM)

1 µL of template DNA (100-500 ng of a complex template – e.g. cDNA, gDNA; 10-50 ng of a simple template – e.g. plasmid DNA)

25 µL of 2 x RedTaq Ready Mix [Sigma]

23 µL of PCR grade water [Sigma]

For some applications e.g. colony PCR the volumes were downscaled five fold.

Following cycling conditions were used:

3.24.1.1. Standard PCR amplification.

94 °C for 1 min

Then 30 cycles of:

94 °C for 1 min

X °C for 2 min, where X is the annealing temperature of the primers ($X=T_m-2$ °C initially, this temperature was tested and further optimized if necessary; melting temperature (T_m) calculated with Oligonucleotide analyzer (M&M: 5.4))

72 °C for 1 to 4 min (depending on the template length usually 1 min per 1 kb)

After completing 30 cycles:

72 °C for 10 min

4 °C hold

3.24.1.2.Touchdown PCR.

Essentially run as described by Don, Cox et al. (1991) and was ran for 5 cycles at each of annealing temperatures until the final (touchdown) temperature was obtained:

94 °C for 1 min

Then 5 cycles at annealing temperature X (calculated as above)

94 °C for 1 min

X °C for 2 min, where X is the annealing temperature of the primer

72 °C for 4 min

Then 5 cycles at annealing temperature X - 1 °C

94 °C for 1 min

X - 1 °C for 2 min, where X is the annealing temperature of the primer

72 °C for 4 min

Repeat 5 times cycle steps until touchdown temperature is reached, then:

72 °C for 10 min

4 °C hold

3.24.1.3.Colony PCR amplification.

Performed essentially as the Standard PCR technique with the following modification: the colony of interest was moved by a sterile toothpick to a PCR tube containing 20 μL of sterile distilled water, resuspended and incubated at 95 °C for 15 minutes. The sample was

subsequently chilled on ice and spun down in the microcentrifuge. 5 μL of the supernatant was used to perform colony PCR reaction adjusting water content in the mix accordingly.

3.24.2. Proofreading polymerase *pfu* Phusion.

Amplification with Phusion [NEB] was performed essentially according to manufacturer's conditions, a typical reaction mixture of 50 μL contained:

0.5 μL of forward primer (100 μM)

0.5 μL of reverse primer (100 μM)

1 μL of template DNA (100-500 ng of a complex template – e.g. cDNA, gDNA 10-50ng of a simple template – e.g. plasmid DNA)

10 μL of 5 X HF Buffer [NEB]

0.5 μL of PCR enhancer 1 U/1 μL [Stratagene] (optional)

0.5 μL of 20 mM dNTPs [Roche]

0.5 μL of *pfu* Phusion polymerase [NEB]

X μL of PCR grade water to 50 μL [Sigma]

Following cycling conditions were used:

3.24.2.1.Standard PCR amplification.

98 °C for 30 s

Then 30 cycles of:

98 °C for 10 s

X °C for 30 s, where X is the annealing temperature of the primer (calculated as above)

72 °C for 1 to 4 min (depending on the template length usually 30 s per 1 kb)

After completing 30 cycles:

72 °C for 10 min

4 °C hold

3.24.2.2. Touchdown PCR.

Essentially run as described by Don, Cox et al. (1991)and was ran for 5 cycles at each of annealing temperatures until the final (touchdown) temperature was obtained:

98 °C for 30 s

Then 5 cycles at annealing temperature X (calculated as above)

98 °C for 10 s

X °C for 30 s, where X is the annealing temperature of the primer

72 °C for 3.30 min

Then 5 cycles at annealing temperature X - 1 °C

98 °C for 10 s

X - 1 °C for 2 min, where X is the annealing temperature of the primer

72 °C for 3.30 min

Repeat 5 times cycle steps until touchdown temperature is reached, then:

72 °C for 10 min

4 °C hold

3.25.Oligonucleotide probe labelling.

3.25.1. 5′ end labelling with γ-^{32}P ATP.

The oligonucleotide probe was 5′ end labeled with γ-^{32}P according to Sambrook (1989) with the following modifications:

Probe labelling for cDNA hybridization (1 blot):

2 μL of T4 polynucleotide kinase [Promega]

2 μL of T4 polynucleotide kinase 10 x buffer [Promega]

1 μL of oligonucleotide mixture (10 pmol)

7 μL of γATP (6000 Ci $mmol^{-1}$; 10mCimL^{-1}) [PerkinElmer]

8 μL ddH_2O

Probe labelling for plaque hybridization (6 membranes):

2 μL of T4 polynucleotide kinase [Promega]

2 μL of T4 polynucleotide kinase 10x buffer [Promega]

1 μL of oligonucleotide mixture (100 pmol)

15 μL of γATP (6000 Ci $mmol^{-1}$; 10mCimL^{-1}) [PerkinElmer]

3.25.2. Random priming labelling with α-^{32}P dCTP.

The PCR amplified hybridisation probe was labelled with the High Prime [Roche] random priming labelling kit using the following procedure: 5.5 μL of DNA probe (30-40 ng) was denatured at 98 °C for 2 min and subsequently chilled on ice for 2 min. After cooling 2 μL of High Prime reagent and 2.5 μL of α-^{32}P dCTP (3000 Ci $mmol^{-1}$) [PerkinElmer] were added. The reaction mixture was incubated at 37 °C for 15 min and the labelling was stopped by an addition of 50 μL of 0.5 M EDTA. The probe was prepared for hybridisation by denaturation at 98 °C for 4 min and subsequent cooling on ice before adding to pre-hybridization solution.

3.26.Alkaline Southern Blotting.

Blotting was performed essentially as described by Southern (1975) and Reed and Mann (1985) using the following buffers and solutions: Pre-hybrid solution (0.25 M potassium phosphate buffer pH 7.2, 7 % (w/v) SDS); Washing buffer 1 (20 mM potassium phosphate buffer pH 7.2, 5 % (w/v) SDS); Washing buffer 2 (20 mM potassium phosphate buffer pH 7.2, 1 % (w/v) SDS); Transfer solution (0.4 M NaOH); 2 x SSC (0.3 M NaCl, 0.03 M Na-citrate, pH 7.0)

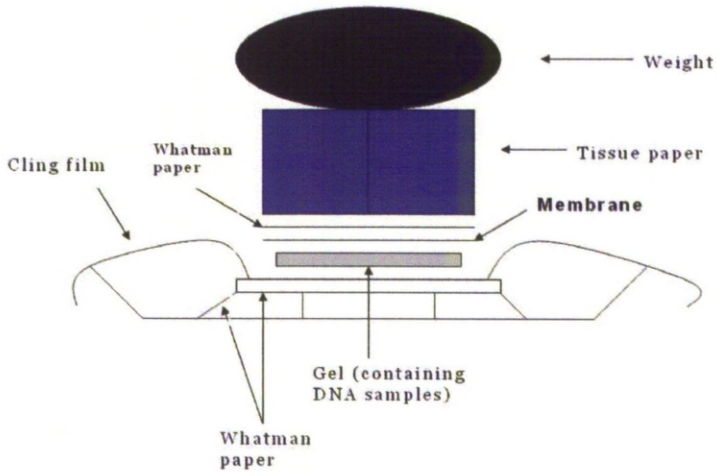


Figure 2.1 Southern blotting apparatus assembly.

All Whatman paper and membrane was pre-soaked in 0.4 M NaOH. DNA digested with restriction enzymes was separated by 1 % (w/v) agarose in TAE. DNA was transferred at room temperature overnight. The membrane was removed and soaked in 2x SSC for 10 min and dried using absorbent paper. The membrane was then left to stand in air at room temperature for 30 min. The membrane was placed onto fresh Whatman paper, and crosslinked with Stratalinker [Stratagene] according to the manufacturer's instructions.

3.27.Colony Blotting.

The Perti dish containing bacterial colonies of interest was moved from 37 °C to 4 °C for at least 2 h prior to the lifting. The appropriately cut and marked nylon membrane [Biorad] was placed onto the plate and allowed to "soak" the colonies for 3 minutes. The membrane was then lifted and placed in a tray on 1 mL of 0.5 M NaOH lift side up for 2 minutes. This step was repeated twice. Subsequent incubation in 1 mL of 1.5 M Tris-HCl pH 7.5 was then used to neutralise the base, this step was repeated twice as well. After neutralisation the membrane was then left to

stand in air at room temperature for 30 min. The membrane was placed onto fresh Whatman paper, and crosslinked with Stratalinker [Stratagene] according to the manufacturer's instructions.

3.28.Hybridization and detection.

The hybridisation work was perfomed in a hybridising chamber Hybridiser HB-1D [Techne]. Following crosslinking, the membrane was subjected to pre-hybridization in 25 mL of pre-hybridization solution containing 0.25 M sodium phosphate buffer pH 7.2, 7 % (w/v) SDS twice for 30 min at 65 °C and subsequently probed with ^{32}P-labelled probe in pre-hybridisation buffer over 48 h at 37 °C. After hybridization it was washed once with 25 mL of washing buffer 1 (20 mM sodium phosphate buffer pH 7.2, 5 % (w/v) SDS) over 30 min at 37 °C and subsequently washed with 25 mL of washing buffer 2 (20 mM sodium phosphate buffer pH 7.2, 1 % (w/v) SDS) for additional 30 min at temperature appropriate for the probe used. Hybridization signal was detected with a phosphor-image screen system [Molecular Dynamics].

3.29.Preparation of cloning adaptors for pYES-MCS expression vector.

The sequence of a 39-bp long polylinker containing EcoRI/XhoI sites was designed. Two complementary oligonucleotides each possessing 3′ A overhangs were designed and synthesized. The two oligonucleotides were mixed and denatured at 95 °C for 15 min and subsequently annealed at 45 °C for 30 min. The product formation was checked by polyacrylamide gel electrophoresis and the insert was TA cloned into the pYES 2.1 expression vector [Invitrogen], restriction analysed to check the orientation of polylinker, confirmed by sequencing and annotated as pYES-MCS

3.30.Cloning and *E.coli* transformation.

PCR products were TA cloned into pGEM T-Easy vector [Promega] following A-tailing (if required) or blunt cloned into one of the following vectors pSCB [Stratagene], pJET 1.2/blunt [Fermentas] according to manufacturers' instructions and transformed using the competent cells provided with the kit or purchased separately.

3.31.Transformation of *S. cerevisiae.*

The transformation of *S. cerevisiae* strains with a URA3 containing vector was performed essentially as described by (Gietz and Schiestl, 2007) using from 250 ng to 1 μg of plasmid DNA per transformation.

4. Protein techniques.

4.1. Protein isolation.

4.1.1. Isolation of extracellular proteins from cultures of *S. aeruginosa.*

Protein was isolated directly from the culture medium. The medium was filtred through Miracloth [Calbiochem] and precipitated overnight with ammonium sulphate (35 % (w/v) at 4 °C) of 5 L of HNC medium. Precipitated protein was pelleted by centrifugation using a Sorvall SLA-3000 rotor at 11,000 g, 4 °C. The pellet was dissolved in 100 mM potassium phosphate buffer pH 6.5 and dialysed overnight against 100 volumes of the same buffer to remove salt.

4.1.2. Isolation of intracellular proteins from *E. coli.*

The *E. coli* cultures were pelleted at 3,000 g and around 2 g of wet biomass was suspended on ice in 2 mL of a suspension buffer (50 mM Tris HCl pH 8.0, 1 mg/mL lysozyme [Fluka], 1 mM PMSF [Sigma]) and sonicated on ice using the amplitude 22 μm and 50 % of power and five cycles of 30 s sonication 30 s break. After sonication the cellular debris was removed by centrifugation at 20,000 g at 4°C and the supernatant was used for downstream applications.

4.2. Protein content assay.

Protein assays were performed according to the method of Bradford (1976) using a calibration curve prepared for bovine serum albumin (BSA) [Bio-Rad]. Appropriate dilutions of protein solution (2 μL) were incubated for 10 minutes at room temperature with 100 μL of Bradford reagent

[Sigma]. After incubation the absorbance was measured at 595 nm with respect to the blank sample containing buffer instead of protein solution.

4.3. Protein activity assays.

4.3.1. Laccase and peroxidase activity assays.

One unit of laccase activity was defined as the amount of enzyme required to oxidise 1 μmol of substrate per minute. Enzyme activity was determined *via* the monitoring of the oxidation of a variety of substrates (Table 2.4) at 42°C for 15 minutes. The assay mixture (1 mL) contained: 890 μL of buffer (differing pH and type of the buffer Table 2.3), 100 μL of 10 mM substrate (dissolved in DMSO) and 10 μL of appropriately diluted enzyme. The molar extinction coefficient of each substrate was used to determine activity (Table 2.1). All measurements were made using a Unicam UV/VIS Spectrometer UV2. The kinetic measurements included variable concentration of the substrate, whereas inhibitory tests included a variable concentration of chosen inhibitor in the buffer and 1 h of preincubation step in 0.1 M potassium phosphate buffer pH 6.5. Thermostability tests also included this preincubation step. Peroxidase activity assays were performed in the same manner with addition of 10 μl of 0.3 % (w/v) H_2O_2 to initiate the reaction.

Table 2.3 Buffers used during various protein work on laccases of *S. aeruginosa.*

pH range	Buffer used	Application
2.0 to 12.0	Britton - Robinson	The effect of pH on activity measurements
2.5 to 7.5	McIlvaine	Routine activity measurements
6.0 to 8.0	Potassium phosphate	Routine storage and purification

Table 2.4 Substrates used during biochemical characterisation of *S. aeruginosa* laccases.

Substrate	Wavelength [nm] λ_{max}	Molar extinction coefficient [cm^{-1} M^{-1}]	Molecular Weight [g/mol]
ABTS	420	36000	548.68
Guaiacol	470	26600	124.14
Veratryl alcohol	310	9300	168.19
Syringaldazine	530	65000	360.37
Dimethoxyphenol	468	49600	154.16
Tetramethylbenzidine	465	37100	240.34
o-dianisidine	460	11300	244.29
Indene	250	9340	116.16
Ferulic acid	312	11000	194.18
Sinapic acid	307	16500	224.21

4.3.2. Haloperoxidase activity assays.

4.3.2.1.Monochlorodimedone (MCD) assay.

Enzymatic activity was measured with the monochlorodimedone assay according to the following protocol derived from Morris and Hager (1966). Haloperoxidase activity was determined at 290 nm *via* the monitoring of the bromination of MCD. The assay mixture (1 mL) contained: 970 µL of potassium phosphate buffer (100 mM, pH 6.5) containing 20 mM KBr and 15 µM MCD and 10 µL of appropriately diluted enzyme. 20 µL of 0.3 % H_2O_2 (w/v) was added to initiate the reaction. The exact initial MCD concentration was estimated spectrophotometrically using an additional blank sample containing all components of the incubation mixture but MCD. The molar extinction coefficient of MCD at 290 nm is 2.02 x 10^4 $M^{-1}cm^{-1}$.

4.3.2.2.Thionin assay.

Enzymatic activity was measured with the thionin assay according to the following protocol derived from Manoj and Hager (2006). Haloperoxidase activity was determined at 598 nm *via* the monitoring of the bromination of thionin [Sigma]. The assay mixture (1 mL) contained: 980 µL of potassium phosphate buffer (100 mM, pH 6.5) containing 10 mM KBr and 25 µM thionin and 10 µL of appropriately diluted enzyme. 10 µL of 0.3 (w/v) % H_2O_2 was added to initiate the reaction. The exact initial thionin concentration was estimated spectrophotometrically using an additional blank sample containing all the components of the incubation mixture but thionin. The molar extinction coefficient of thionin at 598 nm is 6.00×10^4 $M^{-1}cm^{-1}$.

4.3.3. Functional screening of phage lambda library.

4.3.3.1.ABTS screening of submerged plaque library.

Around 5×10^5 pfu were plated on 15 cm Petri dishes of NYZ medium in a form of liquid culture in NZY Top agar according to the manufacturer's instructions. NZY Top agar was supplemented with IPTG to a final concentration of 20 mM and ABTS to the final concentration of 1 mM. Following incubation for approximately 18 h at 37 °C the plate was observed for colour formation

4.3.3.2.ABTS screening of *E. coli* colony lift library.

Around 5×10^5 cfu from phagemid library were plated on 15 cm Petri dishes of LB – amp medium supplemented with IPTG to a final concentration of 20 mM. Following incubation for approximately 18 hours at 37 °C the library was screened in two ways:

- After the incubation the plates were overlaid with 10 mL of 1 mM ABTS solution in McIlvaine buffer pH 3.0 supplemented with 0.8 % agarose and incubated overnight at 37 °C.
- After the incubation the plates were moved to 4 °C for an additional 2 hours and the colonies were lifted from the plates by nitrocellulose membranes. The nitrocellulose membranes were "sandwiched" between filter paper soaked with 1 mM ABTS solution in McIlvaine buffer, pH 3.0, wrapped in cling film to prevent evaporation and incubated overnight at 37 °C.

4.3.3.3. ABTS screening of *S.cerevisiae* library.

Around 5 x 10^5 cfu from *S. cerevisiae* library were plated on 15 cm Petri dishes of SD-U medium supplemented with 2 % (w/v) glucose. Following incubation for approximately 3 days at 30 °C the plates were replica plated with "velvets" onto SD-U medium supplemented with 2 % galactose and allowed to grow for an additional 2 days at 30 °C. The library was screened in two ways:

- After the incubation the plates were overlaid with 10 mL of a 1 mM ABTS solution in McIlvaine buffer pH 3.0 supplemented with 0.8 % agarose and incubated for 24 h at 30 °C.
- After the incubation the plates were moved to 4 °C for additional 2 hours and the colonies were lifted from the plates by nitrocellulose membranes. The nitrocellulose membranes were "sandwiched" between filter paper soaked with 1 mM ABTS solution in McIlvaine buffer pH 3.0, wrapped in cling film to prevent evaporation and incubated for 24 h at 30 °C.

4.3.4. Decolourisation of synthetic dyes with *S. aeruginosa* laccases (adapted from Catherine Ann Houghton's project report).

Decolourisation was determined by monitoring the oxidation of different dyes at 42°C at hourly intervals over a 24 h period. The dilutions of dyes required to obtain an absorbance of 0.8-1.2 were made. The assay mixture contained 990 µL of appropriately diluted dye in Britton-Robinson buffer pH 3.0 and 0.1 ABTS Unit (10 µL) of enzyme. Percentage decolourisation was determined.

4.4. Protein electrophoresis.

4.4.1. One dimensional protein electrophoresis.

4.4.1.1.Denaturing protein electrophoresis.

SDS-PAGE was performed according to the method of Laemmli (1970) using a 12 % stacking gel (37.5:1 acrylamide : bisacrylamide ratio) and 5 % resolving gel at 150 V at room temperature using Mini Protean 3 system [Bio-Rad]. The samples were run until the bromophenol dye reached the end of the gel. Protein bands were visualised with GelCode Blue Safe [Thermo Fisher] unless stated otherwise. PageRuler Plus [Fermentas] was used as a molecular weight marker (typical image of ladder is shown in Appendix 2).

4.4.1.2.Native protein electrophoresis.

Native PAGE was carried out according to the method Ornstein and Davis (Ornstein, 1964) using a 4 % stacking gel (37.5:1 acrylamide : bisacrylamide ratio) and 9 % resolving gel at 4 °C, 100 V. using Mini Protean 3 system [Bio-Rad]. The samples were run until the bromophenol

dye reached the end of the gel. Protein bands were visualised with GelCode Blue Safe [Thermo Fisher] or one of the zymograms.

4.4.2. Two dimensional semi native protein electrophoresis.

4.4.2.1.Preparation of the IPG strips and first dimension separation – isoelectrofocusing.

Non-linear DryStrips [Bio-Rad] were rehydrated overnight in 135 µL of 0.2 % (w/v) CHAPS (Sigma), 20 % (w/v) glycerol, 0.5 % (v/v) Ampholines [Bio-Rad] solution (Hird et al., 2000) containing 25 - 50 µg of protein under mineral oil and focused at following conditions using a Protean IEF Cell [Bio-Rad]:

Ramp	Linear
Temp	20 °C
Rehydration	active at 50 V for 12 h
Pause after rehydration	no
Step 1	250 V for 15 min (linear ramp)
Step 2	4000 V for 2 h
Step 3	4000 V to 20,000 V h^{-1} (rapid ramp)
Step 4	500 V hold

After focusing, one set of strips was stained for activity and another set subjected to second dimension separation using Mini-Protean 3 [Bio-Rad] at constant voltage of 120 V.

4.4.2.2.Second dimension separation.

IPG strips were cleared of mineral oil on absorbent paper and equilibrated in equilibration buffer 1: 35 % (v/v) glycerol 35 % (w/v) urea 2 % (w/v) SDS in 50 mM Tris-HCl, pH 6.8, 2 % DTT for 15 min gel side up and subsequently reequilibrated in equilibration buffer 2: 35 % (v/v) glycerol 35 % (w/v) urea, 2 % (w/v) SDS in 50 mM Tris-HCl, pH 6.8, 2.5 % (w/v) iodoacetamide for another 15 min before placing it on the top of a polyacrylamide gel (5 % stacking gel, 10 % resolving gel). The strip was overlaid with molten agarose and electrophoresed at 100 V at room temperature using Mini-Protean 3 [Bio-Rad]. Protein spots were visualised with GelCode Blue Safe [Thermo Fisher]. PageRuler Plus [Fermentas] was used as a molecular weight marker.

4.5. Western blotting.

The transfer of proteins from unstained SDS-PAGE gels onto Amersham Hybond ECL nitrocellulose membrane [GE Healthcare] was performed essentially as reported by Burnette (1981) using Mini Trans-Blot Cell [Biorad]. After blotting the efficiency of transfer was checked with 0.1 % (w/v) Ponceau S [Sigma] in 5 % acetic acid.

4.6. Immunodetection of 6xHis tagged proteins.

Proteins tagged with 6xHis tag were immuno detected on a nitrocellulose membrane with Monoclonal Anti-polyhistidine Peroxidase conjugated antibodies [Sigma]. After electro transfer the membrane was blocked in 15 mL of PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM NaH_2PO_4, pH 7.4) containing 0.5 % (w/v) BSA [Sigma] from 30 min at room temperature to overnight at 4 °C. The antibody was bound to the antigen during 1 hour incubation of a 1:10,000 dilution of an antibody

solution in PBS containing 0.5 % (w/v) BSA [Sigma] and 0.5 % (v/v) Tween 20 [Sigma]. The excess of antibody was washed off with PBS containing 0.5 % (w/v) BSA [Sigma] and 0.5 % (v/v) Tween 20 with 8 x 5 min washes followed by 4 x 15 min washes. The detection was performed with Amersham ECL Direct [GE Healthcare] according to the manufacturer's instructions. The visualisation of the signal was carried out with Amersham Hyperfilm ECL [GE Healthcare] film and developed with Kodak fixer and developer for autoradiography films [Sigma].

4.7. Zymogram - detection of a laccase activity within the native gel.

Post electrophoresis, the gels (or IEF strips) were soaked in a solution containing 1 mM *o*-dianisidine in 0.1 M McIlvaine buffer pH 3.0. The gels were then gently shaken for 30 min. In the presence of a laccase,*o*-dianisidine is oxidised and appears as brown bands within the gel. Methodology derived from Antipov (Antipov et al., 2003).

4.8. Glycoprotein detection within PAGE.

Detection of glycoproteins was performed using Pro-Q Emerald 488 Glycoprotein Stain Kit [Invitrogen] according to the manufacturer's instructions.

4.9. Protein analysis by tandem mass spectrometry.

4.9.1. In gel tryptic digest of protein samples.

The protein spots from 2D gels were excised and the plugs were destained twice (50 µL/plug) in 50 % acetonitrile in 50 mM ammonium bicarbonate at 37 °C for 15 min. After destaining the plugs were dehydrated with acetonitrile until opaque (10 µL/plug) at 37 °C for 45 min.

After incubation the remaining acetonitrile was removed and the plugs were vacuum dried [Flowgen] for further 30 min at 30 °C. Plugs were rehydrated with 10 µL of 50 mM ammonium bicarbonate containing trypsin (20 ng/ µL) and 2 M urea. The digest was performed overnight at 30 °C. The reaction was stopped with formic acid (1% (w/v) final concentration).

4.9.2. Peptide sequencing.

The samples were run on a tandem mass spectrometer by Mr Mark Prescott from University of Liverpool proteomics facility; the data were sequenced manually using the PepSeq module of MassLynx package [Waters]

4.10.Recombinant protein production.

4.10.1. Production in *E. coli* (BL21 DE3) pET M-11 vector system (adapted from Mohamed Mohamed Noor project report)

Protein expression was induced with 1 mM final concentration of IPTG. Briefly, fresh LB medium with chloramphenicol and kanamycin was inoculated with overnight cultures at 9:1 ratio. Cultures were then grown at 37 °C until midlog phase ($OD_{600} \approx 0.5 - 0.7$) before being induced with IPTG to final concentration 1 mM, with addition of glucose at a final concentration of 1 % (w/v) to increase T7 polymerase-directed gene expression (Novy and Morris, 2001). Seven hours after induction, samples were collected every hour and centrifuged at 2,800 *g* for 15 min at 4 °C and disrupted by sonication.

4.11.Protein purification.

The majority of protein purification experiments were performed with DuoFlow Bio Logic system [Bio-Rad].

4.11.1. Purification of *S. aeruginosa* laccases.

4.11.1.1.Ion exchange chromatography on Q Sepharose.

Around 100 mg of protein solution was mixed with 1/5 (v/v) of Q Sepharose [Sigma] chromatographic medium equilibrated with 100 mM potassium phosphate buffer pH 6.5. The mixture was incubated on roller mixer for an hour. Supernatant containing non bound protein was recovered by centrifugation and bound proteins were eluted with 0.5 M NaCl in 100 mM potassium phosphate buffer pH 6.5 and recovered by centrifugation.

4.11.1.2.Affinity chromatography on ConA Sepharose.

Around 20 mg of diluted protein was loaded onto the chromatographic column packed with 1 mL of ConA Sepharose equilibrated with 100 mM potassium phosphate buffer pH 6.5 containing 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM $MnCl_2$ *via* 50 mL Superloop (GE Healthcare). After sample application the column was washed until no protein was detected in the eluate. Bound proteins were eluted with α-D-methylmannopyranoside (100 mM) until the concentration of protein in the buffer reached zero. Fractions of 1 mL were collected and flow rate of 0.5 mL min^{-1} was applied during whole separation process.

4.11.1.3. Hydrophobic interaction chromatography on Phenyl Toyopearl.

.

Around 10 mg of the ConA eluate was saturated with ammonium sulphate to a final concentration of 0.8 M. Precipitated proteins were separated *via* centrifugation at 20 000 g, 20 °C and discarded. Supernatant was diluted to 50 mL with 0.8 M $(NH_4)_2SO_4$ and loaded onto the chromatographic column packed with 1 mL of Phenyl Toyopearl [Tosoh] equilibrated with 100 mM potassium phosphate buffer pH 6.5 containing 0.8 M $(NH_4)_2SO_4$ *via* 50 mL Superloop (GE Healthcare). After sample application, the column was washed until no protein was detected in the eluate. Bound proteins were eluted with a decreasing step gradient of ammonium sulphate concentration (0.28 M, 0.12 M, 0.064 M, 0 M) in 100 mM potassium phosphate buffer pH 6.5 until the concentration of protein in the buffer reached zero. Fractions of 1 mL were collected and a flow rate of 0.5 mL min^{-1} was applied during whole separation process.

4.11.1.4.Ion exchange chromatography on MonoQ.

Around 0.5 mg of non bound protein sample from HIC concentrated with spin concentrators of 30 kDa cut-off [Millipore] and buffer exchanged to 100 mM Tris HCl pH 9.0 was loaded onto the prepacked MonoQ [GE Healthcare] chromatographic column of 2 mL equilibrated with 100 mM Tris HCl pH 9.0. After sample application the column was washed until no protein was detected in the eluate. Bound proteins were eluted with increasing step gradient of NaCl concentration (0.2 M, 0.4 M, 0.6 M, 0.8 M, 1M, 2M) in Tris HCl pH 9.0 until concentration of protein in the buffer reached zero. Fractions of 0.5 mL were collected and the flow rate of 0.50 mL min^{-1} was applied during whole separation process.

4.11.1.5.Size exclusion chromatography.

Around 0.5 mg of protein concentrated with spin concentrators of 30 kDa cut-off [Millipore] to volume 100 µL were loaded onto the prepacked chromatographic Superdex 75 [GE Healthcare] column of 24 mL equilibrated with 100 mM potassium phosphate buffer pH 6.5 containing 0.15 M NaCl. Separation continued using the same buffer until the concentration of protein in the buffer reached zero. Fractions of 0.5 mL were collected and the flow rate of 0.25 mL min^{-1} was applied during whole separation process.

For estimation of molecular weight of proteins the column was calibrated with Gel Filtration Molecular Weight Markers 12,000-200,000 [Sigma] according to manufacturer's protocol.

5. Bioinformatic software used.

5.1. Sequence analysis and alignments.

All alignments and the assembly of contigs from the isolation of genes by inverse PCR were performed with BioEdit http://www.mbio.ncsu.edu/bioedit/bioedit.html

Plasmid maps and restriction digest predictions were done with VectorNTI Advance 10 [Invitrogen]

5.2. Databases.

- The *A. nidulans* database. http://www.broad.mit.edu/annotation/genome/aspergillus_nidulans/MultiHome.htmL
- The fungal genomics database. https://fungalgenomics.concordia.ca/home/index.php
- National Center for Biotechnology Information database http://www.ncbi.nlm.nih.gov/
- Protein databank http://www.rcsb.org/pdb/home/home.do
- SCOPUS – bibliography search http://www.scopus.com/home.url
- Web of Science – bibliography search http://apps.isiknowledge.com/WOS_GeneralSearch_input.do?product=WOS&search_mode=GeneralSearch&SID=Y22dPOoM2ah946lg4CP&preferencesSaved

5.3. Online tools.

- Protein sequence analysis tools (ProtParam)
  http://www.expasy.ch/
- Sequence manipulation (Sequence Manipulation Suite)
  http://www.ualberta.ca/~stothard/javascript/index.htmL
- Phylogenetic analysis with ClustalW2
  http://www.ebi.ac.uk/Tools/clustalw2/index.html

5.4. Primer design.

PCR Primers, hybridisation probes and cloning adapters were designed with the combination of:

- Primer3
  http://frodo.wi.mit.edu/cgi-bin/primer3/primer3.cgi
- Oligonucleotide analyser
  http://eu.idtdna.com/analyzer/applications/oligoanalyzer/default.aspx

5.5. The 454 Data analysis.

5.5.1. Local database creation.

Contiguous sequences from the transcriptome sequencing of *S. aeruginosa* and *Geomyces sp. P7* cDNAs provided by the Centre for Genome Research along with raw sequence outputs were used to create two local databases for each organism (one with contigs, one with raw data) using BioEdit.

5.5.2. Local BLAST analysis.

The contiguous sequences were analysed for homology with all the non-redundant protein sequences deposited in the NCBI database. The first step of the analysis was downloading the pre formatted non-redundant protein database contained in nr.*tar.gz files from ftp://ftp.ncbi.nih.gov/blast/db/ using CoffeCupFTP program set to binary mode. Once the database was unpacked with gzip the local BLASTX algorithm was run from the command line of Windows using the following syntax:

```
blastall -p blastx -d nr -i [input.fna] -o [output.out]
```

where; [input.fna] – FASTA file containing all contigs, [output.out] – the output text file containing BLAST results in text mode

5.6. Homology modelling of molecular structure.

The molecular structure modelling was performed using FRankenstein Monster approach essentially described by Kosiński, Cymerman et al. (2003). The summary of tools used during the modelling is listed below:

- Modelling and model analysis server https://genesilico.pl/toolkit/
- Modeller – comparative protein modelling program http://salilab.org/modeller/download_installation.htmL
- SwissProt – comparative protein modelling program http://swissmodel.expasy.org/
- Phyre – fold recognition server http://www.sbg.bio.ic.ac.uk/~phyre/
- Verify3D – model quality assesment http://nihserver.mbi.ucla.edu/Verify_3D/

Chapter 3

THE PURSUIT OF A NOVEL FUNGAL HALOPEROXIDASE

1. Introduction.

This project originated as the continuation of Jonathan K. Moore's doctoral work; "Pursuit of a novel haloperoxidase for asymmetric biocatalysis" (Moore, 2007) and at its initial stages were directed according to guidelines drawn by Moore i.e. selection of organisms and culturing methods, protein isolation methods and conditions of enzyme activity assays.

In his doctoral work, Moore worked on two proteins. Initially the work was carried out on putative chloroperoxidase of a basidiomycete fungus *Agaricus bisporus* that was found with the BLAST search of fungal homologs of the best known chloroperoxidase (CPO) from *Caldariomyces fumago*. Unfortunately this approach did not result in the isolation of the gene and the very low levels of enzymatic activity in the native extracts suggested that further work on this project would be unproductive (Moore, 2007). Another approach involved an identification of a haloperoxidase from a basidiomycete strain. Moore screened two known organohalogen producers: *Stropharia aeruginosa* CBS 839.87 and *Bjerkandera fumosa* CBS 152.79 (Verhagen et al., 1996). On the basis of activity assays with monochlorodimedone (MCD) on extracellular extracts it was identified that *Stropharia aeruginosa* secretes a CPO-like protein.

Two strategies that had previously been suggested by Moore were used to try to find a novel fungal haloperoxidase. Parallel attempts were made to a) follow Moore's work on *S. aeruginosa* putative haloperoxidase

and b) to identify a novel haloperoxidase through BLAST analysis of proteins sequence similar to CPO.

2. *Aspergillus nidulans* as a source of putative fungal haloperoxidase.

As part of the pursuit of a novel fungal haloperoxidise, genome mining using BLAST searches was used to identify homologs of the *Caldariomyces fumago* chloroperoxidase. Sixteen of the most similar sequences to that of this well known halogenating enzyme were selected (with E value scores below 1e-05) and multiply aligned using the Clustal W module of BioEdit program (Figure 3.1). One of the most similar sequences was a putative chloroperoxidase of *Agaricus bisporus*; however the previous work carried out by Moore was unsuccessful. Therefore the attention was drawn to the second highest scoring hypothetical protein: *Aspergillus nidulans* AN7823, another putative haloperoxidase. According to BLAST, this protein possesses one of the highest sequence similarities to the *C. fumago* enzyme, with 40 % sequence similarity, 29 % sequence identity and 100 % conservation of the haem-binding site (Figure 3.1), making it a promising target for further studies. A literature search and bioinformatics analysis of the sequence was conducted to obtain a more detailed picture of this putative haloperoxidase.

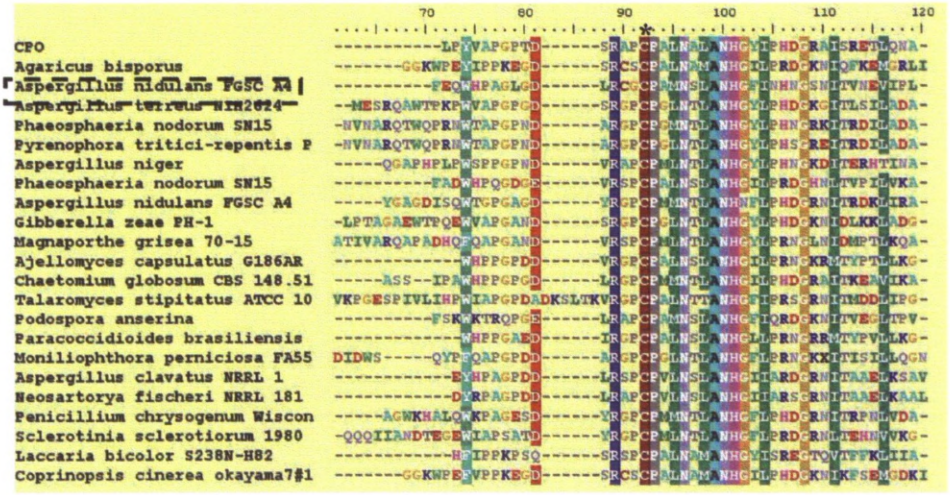


Figure 3.1 Multiple Sequence Alignment of CPO showing its fungal homologs.

Multiple Sequence Alignment (MSA) of the *C. fumago* chloroperoxidase heme binding region with proteins showing the highest sequence similarity, identified by BLAST searching. The conserved proximal haem binding residue (Cys) is marked with an asterisk (*). Identical residues are shaded with the same colour. The full protein alignment is presented in Appendix 1.

3. Further bioinformatics analysis of AN7823 protein.

Searching the *A. nidulans* genome database of the BROAD institute (Galagan et al., 2005), it was found that AN7823 protein is encoded by a single 918 bp exon that translates into a 305 amino acid product belonging to the peroxidase II family of proteins (Figure 3.2). This family includes secreted fungal peroxidases like lignin peroxidases and manganese-dependent peroxidases. The gene is situated on the + strand of chromosome IV of *A. nidulans,* within the sterigmatocystin synthesis gene cluster (Brown et al., 1996). Sterigmatocystin is a polyketide, a secondary fungal metabolite known to be very toxic and mutagenic. Brown et al. annotated the protein as putative chloroperoxidase StcC but no further

published work in the area of biocatalysis has been carried out on the gene or its product.

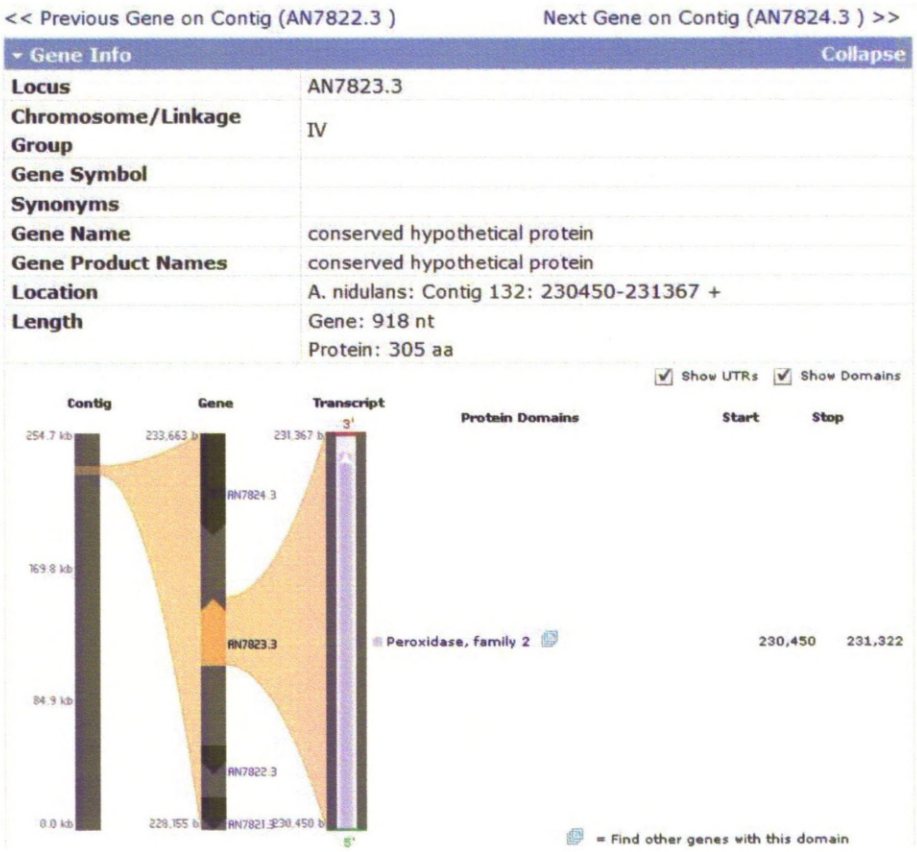


Figure 3.2 Details of the AN7823 gene and its positioning from the BROAD Institute database of *A. nidulans*.

AN7823 is a 918 nt open reading frame located on contig 132. AN7823 is situated on the + strand of chromosome IV of *A. nidulans* within the sterigmatocystin synthesis gene cluster. It is situated in reverse orientation to the surrounding P450 monooxygenase gene (AN7824) and an unknown gene (AN7822) from the sterigmatocystin synthesis gene cluster.
The AN7823 ORF encodes a putative protein product of 305 amino acids.

The sterigmatocystin synthesis cluster is composed of 20 different genes positioned in very close proximity on chromosome IV. The cluster also encodes an array of other enzymes, including a polyketide synthase, fatty acid synthases, monooxygenases, dehydrogenases, an esterase and a reductase, most of them of unknown function in sterigmatocystin biosynthesis. In particular, a function has not yet been assigned to the StcC enzyme in the proposed biosynthetic pathway (Brown et al., 1996). The StcC nucleotide sequence was translated into amino acid sequence and analysed in more detailed manner with Pfam (Finn et al., 2008) to search for similarities to known domains. This was followed by a PeroxiScan search (Koua et al., 2008) to precisely assign AN7823 to the correct group of peroxidases. Signal IP (Bendtsen et al., 2004) was then used to detect signal sequences that are used in the secretion of the polypeptide. Protparam (Gasteiger et al., 2005) was used to calculate the basic parameters of the protein product such as molecular weight, PI and stability index and NetNGlyc was used to find putative N glycosylation sites (Gupta et al., 2004). The Pfam domain search suggested that the protein indeed belongs to family II of the peroxidases which are mainly secretory monomeric proteins. A more detailed prediction performed with PeroxiScan (Koua et al., 2008) showed that the AN7823 has the motif of haem haloperoxidase within the family II of peroxidases. The signal peptide cleavage site by signal peptidase I was identified by SignalIP 3.0 using both neural networks (NN) and hidden Markov models (HMM) at position 25-26. Predictions of N-glycosylation sites were performed using NetNGlyc which identified 3 highly probable N glycosylation sites at the positions 59, 62 and 201. The prediction performed with ProtParam showed that the protein, after signal peptide cleavage, should be around 30 kDa with the theoretical pI of 5.2 and quite unstable characteristics with an instability index of 44.4, where a protein is defined as a stable one if its instability index is less than 40.

MLLKSIQNIVCGLVPTFFLFGSAAAELDFEQWHPAGLGDLRCGCPAMNSL
ANHGFINHNGSNITVNEVIPLMQEVFHLSEELATIVTGLAVLSADDPASG
IFNLDMLNRHNIFEHDASLTRKDFYLGGDGHTIDQPTLDEFLSYFDGKEW
IDLNDAAAARYARVLDSREKNPSFLYQDQQLITSYGETIKYFRTMVDPRS
NKTSAEFVRILFTEERLPVREGWQRPREEISGFSLASDVVQLALRTPEKF
IGMPFDQRPFAEQAFDPLPWQRPPIWTPPNYPGFSKRHFSELVGRFAKKA
LPFRA*

Figure 3.3 Summary of putative *A. nidulans* CPO-like protein sequence analysis findings.

Summary of bioinformatic findings on AN7823 CPO-like protein Signal peptide, signal peptide cleavage site AAA/EL, N-glycosylation sites, C-thiolate proximal heme ligand, N, R, N – stabilization of heme *via* hydrogen bonding Function annotated using CPO crystal structure (Sundaramoorthy et al., 1995).

4. Amplification and cloning of AN7823 gene.

Although there was bioinformatic evidence that *A. nidulans* encodes a potential haloperoxidase, there was no direct biochemical evidence to confirm these observations. Therefore, an *o*-dianisidine assay was performed using extracellular extracts of *A.nidulans WT bi A1* grown under sterigmatocystin synthesis-inducing conditions (Brown et al., 1996). These extracts tested positive for peroxidase activity by the *o*-dianisidine assay (data not shown, McBride, J. – project report) and it was therefore decided to express AN7823 protein in a heterologous host. As the gene has no introns, the coding sequence could be amplified directly from genomic DNA. Genomic DNA of *A. nidulans* WT *bi A1* was extracted and used as a template for PCR amplification. Two sets of primers were designed to

amplify the gene with and without the sequence encoding the signal peptide (Figure 3.4); these were designated SP and NS respectively. Both PCR products (Figure 3.5) were A-tailed and cloned into the pGEM T-easy vector. Sequencing confirmed that the cloned sequences were correct and the clones were designated AN7823-SP and AN7823-NS respectively.

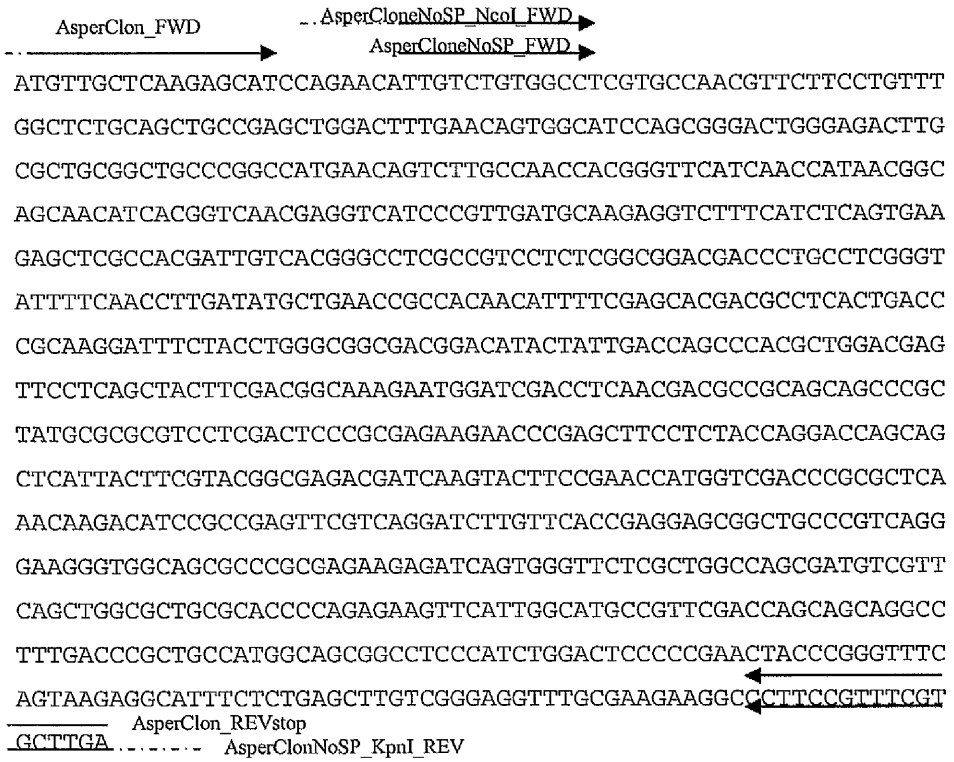


Figure 3.4 Summary of primer binding sites that were used to amplify the AN7823 gene for cloning into various vectors.

AsperClon_FWD – Amplification from start codon
AsperCloneNoSP_FWD – Amplification without signal peptide start codon added to primer
AsperCloneNoSP_NcoI_FWD – Amplification without signal peptide with start codon added to primer in NcoI restriction site
AsperClon_REVstop – Amplification with stop codon
AsperClonNoSP_KpnI_REV Amplification with stop codon and KpnI restriction site

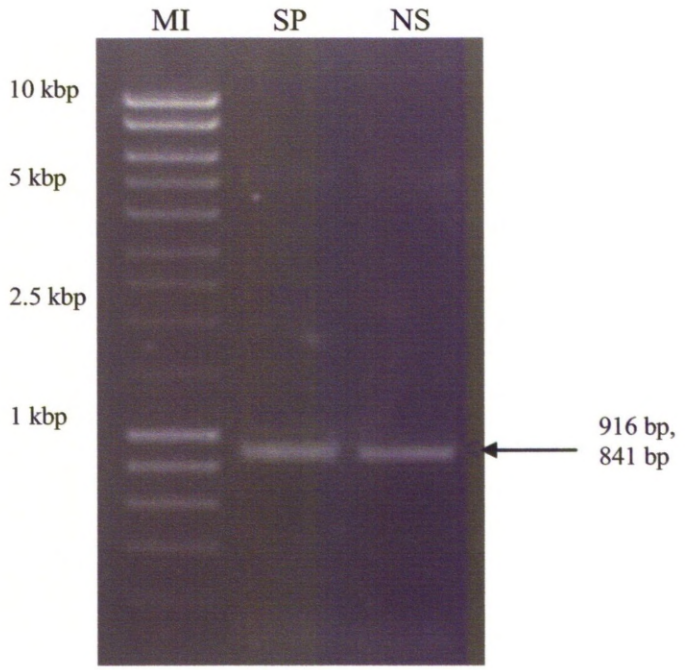


Figure 3.5 PCR amplification of the AN7823 gene from genomic DNA of *A. nidulans WT biA1*.

PCR amplification of the full length sequence was carried out with a primer pair AsperClon_FWD, AsperClon_REVstop. Truncated sequence (without the sequence encoding signal peptide) was amplified with a primer pair AsperCloneNoSP_FWD, AsperClon_REVstop (Figure 3.4). MI – molecular weight marker Hyperladder I; SP – full length amplification of AN7823; NS – truncated amplification of AN7823. Electrophoretic separation was performed on 1 % (w/v) agarose gel.

5. Expression of AN7823 protein in *E. coli.*

5.1. Preparation of construct.

Once the AN7823 gene had been cloned into the cloning vector an attempt was made to express an N-terminally His-tagged version of the protein lacking the signal sequence using an IPTG-inducible *E. coli* expression system. This required subcloning of the gene into the pET-M11 vector (Dümmler et al., 2005) which was carried out by the introduction of an NcoI restriction site at the ATG initiation codon, in order to create the N-terminal 6xHis tag-Tev fusion protein. The NcoI site was introduced by amplifying the gene with a pair of primers containing restriction sites NcoI/KpnI from the AN7823-NS construct using proofreading polymerase. The resulting PCR product was A-tailed and cloned into the pGEM T-easy vector according to the manufacturer's instructions. The DNA of a new clone was confirmed by sequencing and annotated as AN7823-NS-NcoI/KpnI (Figure 3.6) As the native sequence already contained two NcoI sites the aim was to perform a KpnI digest of the AN7823-NS-NcoI/KpnI clone followed by partial NcoI digest in order to generate the correct fragment for subcloning. The expected size of the AN7823 flanked by restriction sites was 844 bp (Figure 3.7); the expected size of digested pET-M11 vector was 5360 bp.

Unfortunately, this attempt was not successful, probably due to inefficient ligation and transformation of competent cells with the construct. In order avoid the use of partial digests and facilitate future subcloning, it was decided to remove both internal NcoI sites. This was performed by introducing two silent mutations using a QuikChange site directed mutagenesis kit with two sets of mutagenic primers, according to the manufacturer's instructions. The mutation sites are shown in Figures 3.8 Figure 3.9 summarise both rounds of mutagenesis. The final construct

was confirmed by sequencing and designated AN7823-NS-NcoI/KpnI-2Nco_Mut. The gene was then subcloned into pET-M11 expression vector using NcoI and KpnI sites. The isolated plasmids were analysed by restriction digest, the sequence was confirmed and annotated as pET-M11-AN7823_2NcoI_Mut.

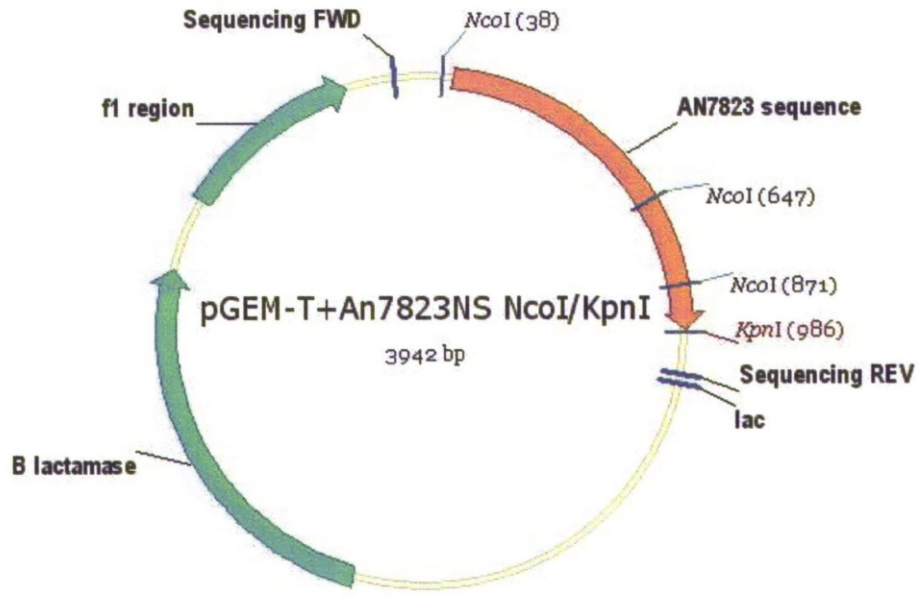


Figure 3.6 Map of the cloning vector pGEM-T with cloned AN7823 sequence amplified with primers introducing cloning sites NcoI and KpnI.

This plasmid was used for gene cloning and subsequent mutagenesis that removed two internal NcoI sites. The signal peptide sequence was removed from AN7823 to allow the N-terminal fusion of this protein with 6xHis tag provided by pET-M11 cloning vector.

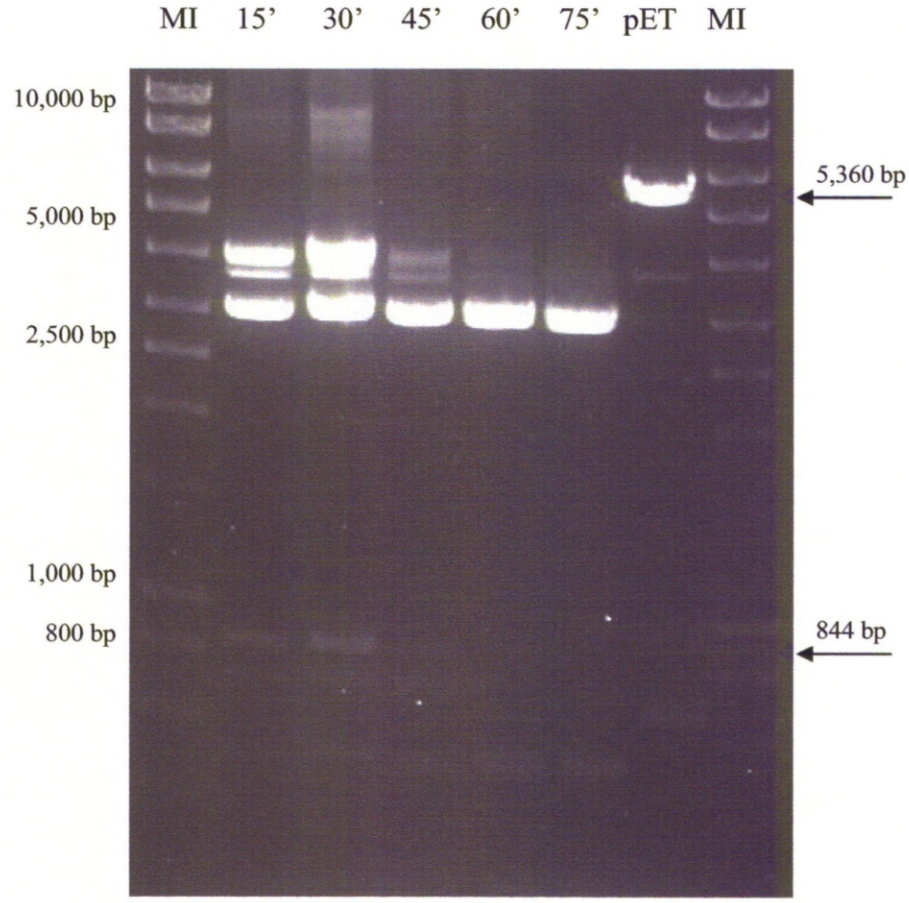


Figure 3.7 Partial digest of linearised with KpnI AN7823NS-NcoI/KpnI construct with NcoI restriction enzyme.

The AN7823NS-NcoI/KpnI plasmid (Figure 3.6) containing the sequence of AN7823 without the signal peptide was linearised with KpnI then subjected to a time course (15 to 75 minutes) digestion with NcoI. The required 844 bp KpnI/NcoI fragment is indicated with an arrow. Lane pET shows pET-M11 expression vector digested with KpnI and NcoI to give a linear fragment of the indicated size. MI - Hyperladder I size markers. Electrophoretic separation was performed on 1 % (w/v) agarose gel.

NS-NcoI/KpnI-2NcoI_Mut.

```
Original sequence:

5` CGATCAAGTACTTCCGAACCATGGTCGACCCGCGC 3`

MutaNcoI_1 fwd
5` CGATCAAGTACTTCCGAACGATGGTCGACCCGCGC 3`
3` GCTAGTTCATGAAGGCTTGCTACCAGCTGGGCGCG 5`
MutaNcoI_1 rev


Original sequence:

    5` CCTTTGACCCGCTGCCATGGCAGCGGCCTCCCATCTGG 3`

MutaNcoI_2 fwd
    5` CCTTTGACCCGCTGCCGTGGCAGCGGCCTCCCATCTGG 3`
3`CCAGATGGGAGGCCGCTGCCACGGCAGCGGGTCAAAGG 5`
MutaNcoI_2 rev
```

Figure 3.8 Site directed mutagenesis of internal NcoI sites of AN7823NS-NcoI/KpnI plasmid vector containing the sequence of AN7823 without the signal peptide, and two cloning sites compatible with pET-M11.

The proposed introduction of silent mutations in two internal NcoI sites (C$^{\downarrow}$CATGG) of AN7823NS-NcoI/KpnI plasmid vector.

Site 1: AN7823-NS-NcoI/KpnI. NcoI site I mutation (C → G) changed codon ACC to ACG both encoding threonine. Site 2: AN7823-NS-NcoI/KpnI. NcoI site II mutation (A → G) changed codon CCA to CCG both encoding proline.

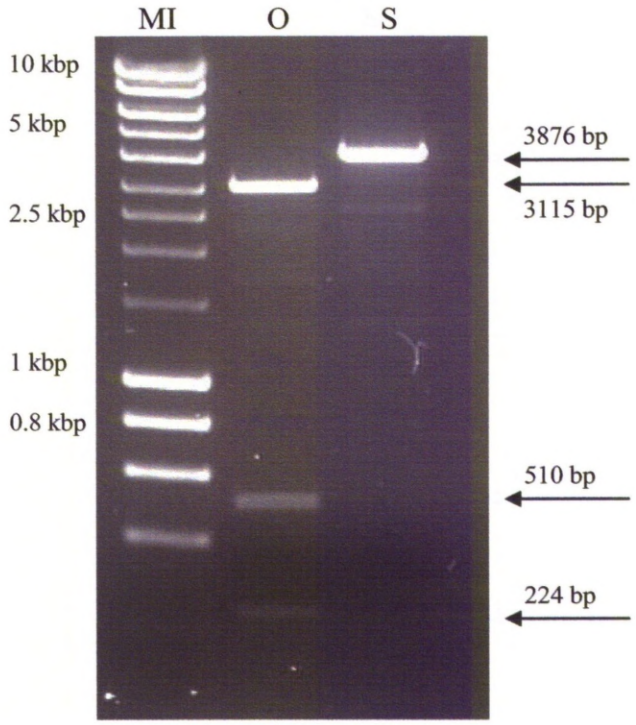


Figure 3.9 NcoI digest of native AN7823-NS-NcoI/KpnI and mutated AN7823-NS-NcoI/KpnI-2Nco_Mut plasmids confirming the removal of two internal NcoI sites to facilitate gene cloning into an expression vector pET-M11.

In order to facilitate sub-cloning of the AN7823 sequences into expression vectors it was necessary to remove internal NcoI sites. This was achieved by site-directed mutagenesis; the plasmids were digested with NcoI to confirm that both internal NcoI sites had been removed. MI – molecular weight marker Hyperladder I; O – original construct AN7823-NS-NcoI/KpnI digested with NcoI; S – mutagenised construct AN7823-NS-NcoI/KpnI-2Nco_Mut digested with NcoI. Electrophoretic separation was performed on 1 % (w/v) agarose gel.

5.2. Expression of AN7823 protein in *E. coli.*

The sequence pET-M11-AN7823_2NcoI_Mut construct was confirmed, map of the construct is presented on Figure 3.10. Attempts were then made to express the 35 kDa His-AN7823 protein in *E. coli* BL 21 (DE3). This work was performed by an undergraduate project student Mohammed Radzi Mohammed Noor.

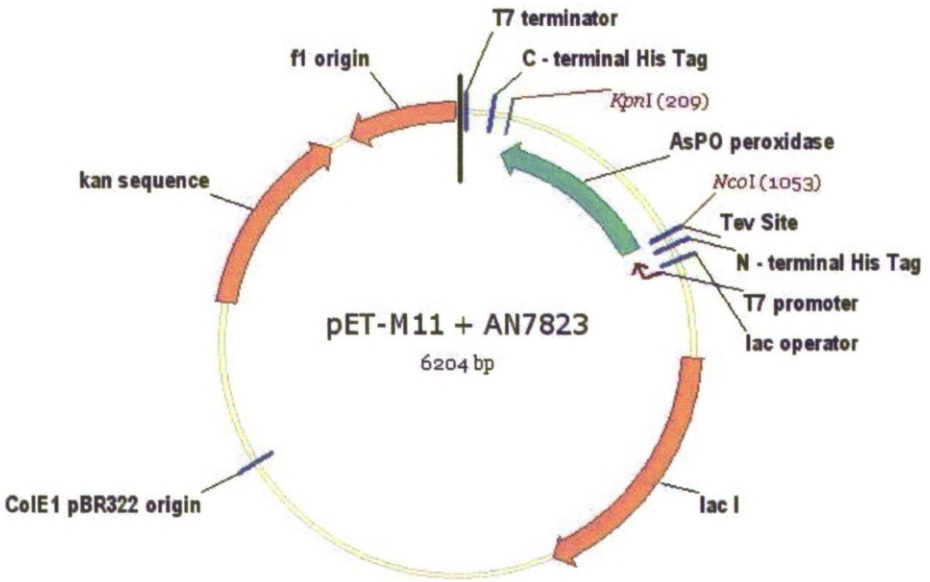


Figure 3.10 Map of the *E. coli* expression plasmid pET-M11-AN7823_2NcoI_Mut.

This plasmid was constructed to express the putative *A. nidulans* AN 7823 peroxidase in *E. coli*. It introduces the N-terminal 6xHis tag to the protein sequence to facilitate protein purification.

A time course experiment was performed to identify the optimum time for expression. Protein expression was induced and analysed using SDS-PAGE and Western blotting. The amount of soluble fusion His-AN7823 protein was minimal and detectable only by Western blotting with anti-6xHis monoclonal antibody (Figure 3.11), with the vast majority of the protein expressed as insoluble inclusion bodies (Figure 3.12). In order to limit the formation of inclusion bodies of the His-AN7823 protein, different strains of *E. coli*, providing tighter control of expression, i.e. BL21 pLysS and BL21 pLysE were tested along with various temperatures and various concentrations of the inducer, however none of these strategies showed any improved expression levels.

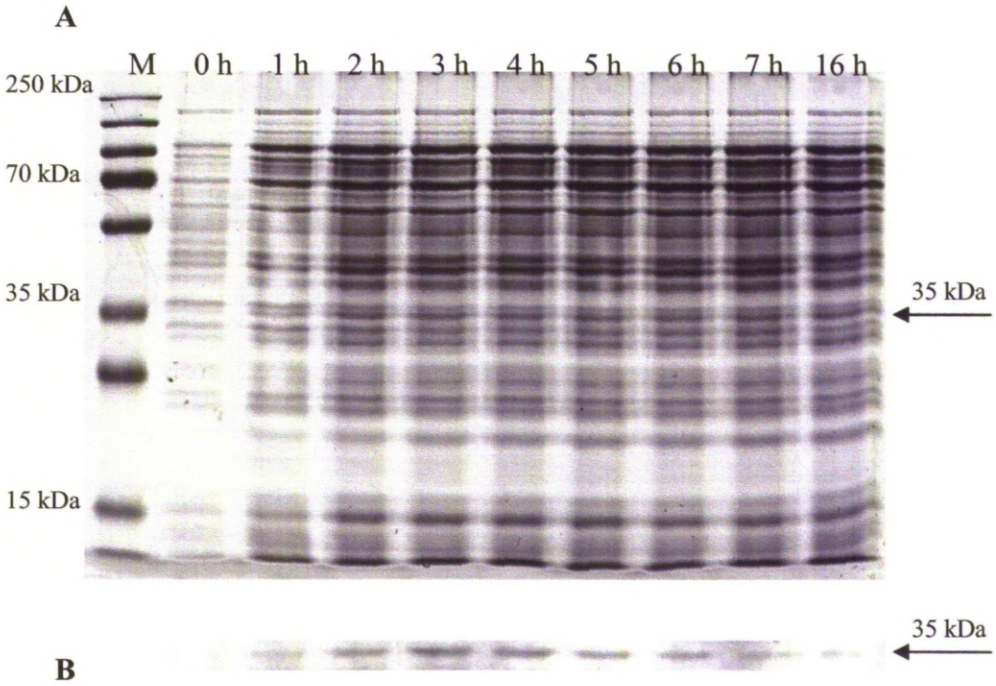


Figure 3.11 SDS-PAGE (A) and Western blot (B) of soluble protein samples from *E. coli* BL21 (DE3) pLysE cells expressing AN7823.

N-terminally 6xHis-tagged *A nidulans* putative peroxidase AN7823 was expressed from plamid pET-M11-AN7823_2NcoI_Mut in *E coli* BL21 (DE3) pLysE. After sonication and centrifugation, samples of soluble

protein were separated on a 12 % SDS polyacrylamide gel and stained with GelCode blue (A). Samples were taken before induction (0 h) and at the indicated times after induction with 1 mM IPTG. M – molecular weight marker, with sizes as indicated. (B) Western blot of the gel in A, probed with anti 6xHis antibody. The blot segment corresponding to the predicted molecular weight of the expressed protein (35 kDa) is shown. The amount of soluble protein is minimal.

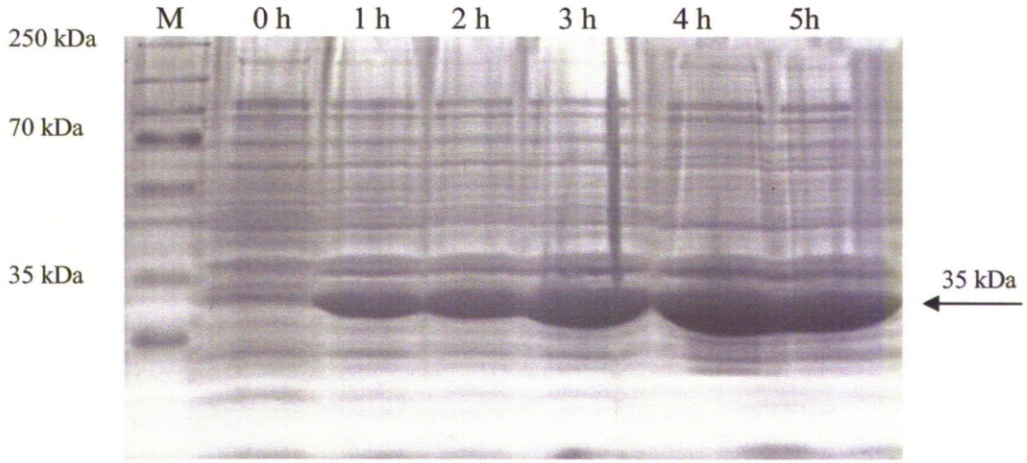


Figure 3.12 SDS-PAGE of the insoluble protein samples from *E. coli* BL21 (DE3) pLysE cells expressing AN7823.

N-terminally 6xHis-tagged *A. nidulans* putative peroxidase AN7823 was expressed from plamid pET-M11-AN7823_2NcoI_Mut in *E coli* BL21 (DE3) pLysE. After sonication and centrifugation, samples of cellular debris were separated on a 12 % SDS polyacrylamide gel and stained with GelCode blue. Samples were taken before induction (0 h) and at the indicated times after induction with 1 mM IPTG. M – molecular weight marker, with sizes as indicated. The vast majority of the protein is present as insoluble fraction (inclusion bodies).

6. Summary of AN7823 protein work.

A putative haloperoxidase of *A. nidulans* was identified *via* BLAST searching for a *C. fumago* homolog. It was found that the protein being a part of sterigmatocystein synthesis cluster is encoded by a single exon and shows similar predicted features to CPO (a secreted haem glycoprotein, showing 100 % conservation of haem binding residues). The gene encoding the protein was amplified without the signal peptide and cloned into the expression vector pET-M11 to create an N-terminally fused 6xHis tagged protein. Expression of the protein was at a very high level; however the vast majority of the protein was in the form of insoluble inclusion bodies. To address this problem several *E. coli* strains and methods of expression were tested in a course of undergraduate project of Mohammed Radzi Mohammed Noor, without much success. The work AN 7823 was abandoned and efforts were then shifted to *S. aeruginosa* proteins.

Chapter 4

ISOLATION, PURIFICATION AND CHARACTERIZATION OF *S. AERUGINOSA* OXIDOREDUCTASES.

1. Introduction.

In a pursuit of novel fungal haloperoxidase, Moore described the identification of a novel enzyme from a basidiomycete strain *Stropharia aeruginosa* CBS 839.87. As an indicator of enzymatic activity a monochlorodimedone (MCD) assay (Hager et al., 1966) of extracellular proteins precipitated with ammonium sulphate was used in a conjunction to an oxidase-specific assays of *o*-dianisidine and tetramethylbenzidine. The principle of the MCD assay is presented in Figure 4.1, whereas *o*-dianisidine assay is described in detail in section 2.

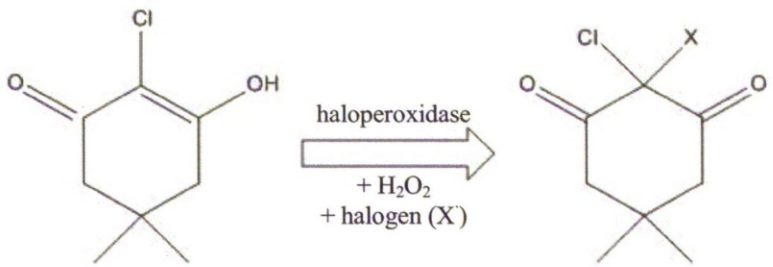


Figure 4.1 The monochlorodimedone (MCD) assay (Hager et al., 1966) that is used to measure haloperoxidase activity.

MCD absorbs at 290 nm. In the presence of hydrogen peroxide and either bromide or chloride ions, haloperoxidases catalyse the bromination or chlorination of MCD. The change can be monitored as a decrease in absorption in a spectrophotometer due to a decline in concentration of the chromogenic substrate.

The MCD assay however has been known to be problematic in that it yields false positives, since it is focused on decline of the substrate not the formation of product (Wagner et al., 2008). Furthermore, the wavelength used in analysis (290 nm) coincides with the long tail of the tryptophan absorbtion spectrum which can contribute to the absorbance at high protein content and low enzyme activity.

Initially, attempts were made to reproduce the results presented by Moore regarding the chlorinating activities of extracellular proteins secreted by *S .aeruginosa*. The fungus was grown and extracellular protein extracts were prepared according to the methods described by Moore (2007). Unfortunately no chlorinating activity was detected in the samples not only with monochlorodimedone assay but also with the more reliable thionin assay (Manoj and Hager, 2006). Both methods were modified in many ways (substrate concentrations, pH and reaction temperatures) but each time the extracts of *S. aeruginosa* failed to show any chlorinating activity.

The conditions of protein isolation described by Moore (overnight precipitation at 45 % (w/v) ammonium sulphate) resulted in a good yield of precipitated protein; however co-precipitation of a dark contaminating pigment, probably melanin (Butler and Day, 1998) (Figure 4.2) was believed to be a significant problem for activity assay measurements due to the need to dilute the protein sample significantly before any reliable absorbance measurements could be performed. It was initially suspected that the sample might be diluted below the detection limit of both haloperoxidase assays, therefore liberating protein samples from the contaminating pigment became a priority before any meaningful measurements of halogenating activity could be performed.

Figure 4.2 Concentrated pigmented sample of *S. aeruginosa* extracellular proteins.

Sample precipitated with ammonium sulphate (35% (w/v) overnight at 4 °C and concentrated using spin concentrators, membrane cut-off 30 kDa.

2. Pigment removal and initial enzyme characterisation.

The aforementioned pigment was a significant problem, not only for activity measurements but also for the protein purification process. It was reported that the pigment binds irreversibly (or very strongly) to anion exchangers such as Q Sepharose "rendering the column useless for further purifications" (Moore, 2007). Initial trials to separate a presumed small molecule of pigment from substantially larger proteins was performed using a BioRad DuoFlow system equipped in Sephadex G75 column, however this resulted in clogging the column due to strong pigment accumulation at the inlet in a form of dark brown sticky "mud", rendering the size exclusion medium useless for any further work and requiring thorough cleaning of the chromatographic system.

In order to liberate the protein sample from the pigment a variety of methods (e.g. ethanol precipitation, polyethylene glycol treatment) was tested and proved ineffective. Therefore an attempt was made to exploit the

observation made by Moore and employ Q Sepharose to strongly bind the pigment followed by liberation of all the proteins using high ionic strength. The process of protein purification started from the optimization of salting out conditions and their effect on the activity of the precipitated protein and the protein content. Due to the lack of a reliable halogenating activity assay, the *o*-dianisidine peroxidise assay (Figure 4.3) was used; as haloperoxidase should also exhibit an ancillary peroxidise activity. Extracellular proteins were precipitated at 4 °C overnight with ammonium sulphate concentrations of 25, 35 and 45 % (w/v) and tested for both protein content and specific activity towards *o*-dianisidine. At the same time the pigment content was checked visually comparing samples of the same protein concentration. Figure 4.4 shows a SDS PAGE gel and native PAGE gel stained with *o*-dianisidine of this optimization process.

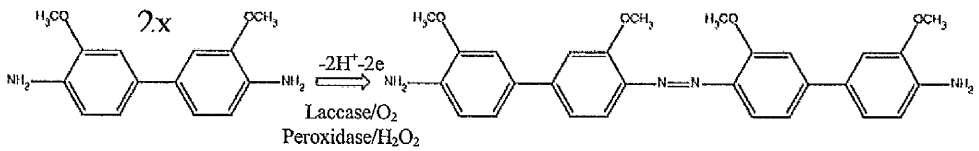


Figure 4.3 Principle of the *o*-dianisidine peroxidase/laccase assay.

In the presence of an oxidation catalyst (peroxidase or laccase) *o*-dianisidine becomes oxidised to form a coloured product bisazobiphenyl through coupling of two quinonediimine molecules. The appearance of this product can be measured spectrophotometrically at 460 nm. The product is weakly soluble in aqueous solution and can be employed as a good method of activity gel staining

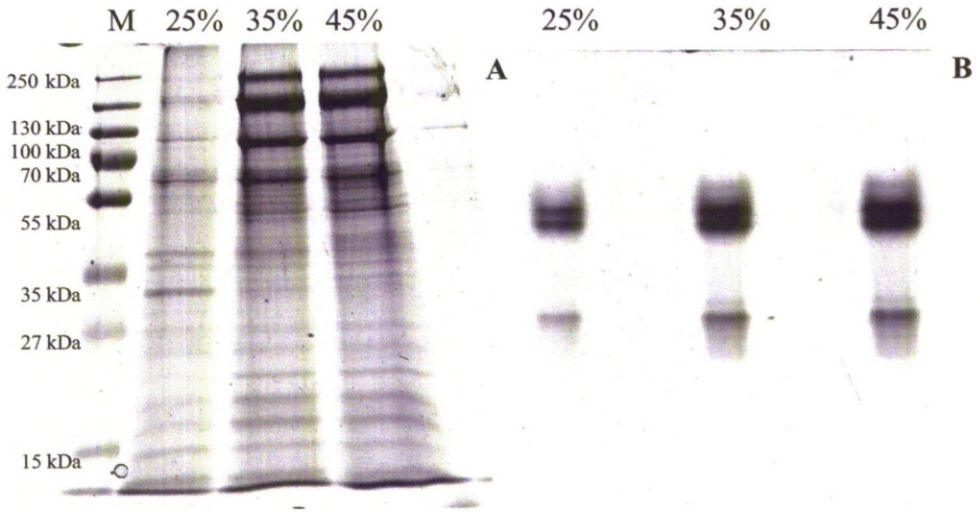


Figure 4.4 Optimisation of the ammonium sulphate precipitation conditions of extracellular protein from *S. aeruginosa* grown on HNC culture medium.

The protein solution was subject to ammonium sulphate precipitation (25, 35 and 45 % (w/v)). Around 10 µg of precipitated protein was then analysed by 12 % SDS-PAGE gel (A) and a duplicate 9 % native gel was stained with *o*-dianisidine (B).

As can be seen on Figure 4.4 no difference in protein yield or activity was observed between ammonium sulphate concentration 45 % (w/v) and 35 % (w/v). As a result the amount of ammonium sulphate used for salting out was decreased from 45 % (w/v) (Moore, 2007) to 35 % (w/v) for overnight precipitation at 4 °C, which significantly decreased the amount of co-precipitated pigment without affecting the amount of precipitated *o*-dianisidine-active protein. The salted out protein was then dialysed against phosphate buffer pH 6.5 to remove the salt and to attempt to trap the pigment on Q Sepharose was performed.

Once the optimal conditions of protein precipitation were found, the binding characteristics of Q sepharose were used to separate the protein sample from the pigment. Instead of using a packed column it was decided to incubate the resin with the pigmented protein sample in batch; on a rotating mixer to ensure uniform binding of the pigment to the

chromatographic medium . Some active protein that was bound to the exchanger was recovered by elution with 0.5 M NaCl with minor pigmentation. The eluted proteins tested negative for haloperoxidase activity, and subsequent attempts to elute more protein from the resin with 2 M NaCl or 0.5 % (v/v) Triton X-100 yielded traces of protein without haloperoxidase activity. Several attempts were made with different elution conditions, and each tested negative for haloperoxidase activity. It was assumed that the halogenating activity reported by Moore was either artifactual, due to the well-known unreliability of MCD assay (Wagner et al., 2008), or that the protein was impossible to separate from the pigment that obscures spectrophotometric readouts of MCD activity. It was claimed by Wagner, Molitor et al. (2008) that other high redox enzymes (cytochrome C family, lignin peroxidases) can cause reactions that result in the disappearance of MCD without the formation of a halochlorodimedone. They found that some lignin peroxidases can directly oxidise MCD to form a product that has a different absorption spectrum to that of MCD causing the maximal absorbance peak of the products to shift from 290 nm to 270 nm and giving a false positive for the MCD assay (Farhangrazi et al., 1992, Wagner et al., 2008). A real positive result in a MCD assay causes the absorbance to disappear, not shift. Moore performed the assay at 290 nm and the decline of absorbance he observed might have been due to MCD conversion by oxidation with *S. aeruginosa* lignin peroxidases. As a result of these findings the attention turned to the enzymes of the ligninolytic cycle which had shown the highest activities (0.6 Units per mg of protein (Moore, 2007)). An ABTS assay that is commonly used for laccases has been since employed as a routine method of activity measurements; the principle of this assay is presented in Figure 4.5. Using this assay the pool of depigmented proteins was confirmed positive for laccase activity

3. Laccase isolation and purification.

In order to find a suitable medium for another step of purification it was decided to try to use the fact that many extracellular fungal laccases are N-glycosylated. To confirm this prediction, an extracellular extract was loaded in duplicate on a native gel, and half of the gel was stained with *o*-dianisidine to detect the activity whilst the other half was stained with glycoprotein staining kit to detect glycoproteins; both gel images are shown in Figure 4.6 and indicate that bands of laccase activity correspond to bands of glycoproteins. It was therefore decided to use a Concanavalin A column as a next chromatographic step. After the previous two procedures (precipitation and Q sepharose chromatography) the sample still contained minor amounts of pigment that no longer affected analysis unless the sample was significantly concentrated. Hence, it was decided to perform the next chromatographic step in large loading volumes using a 50 ml dynamic loop. ConA-agarose medium was shown to bind the proteins of interest and it did not show electrostatic characteristics similar to Q Sepharose, this prevented the pigment from binding to the resin and allowed pigment-free sample concentration during the elution with high concentrations of α-D-methylmannopyranoside (Figure 4.8).

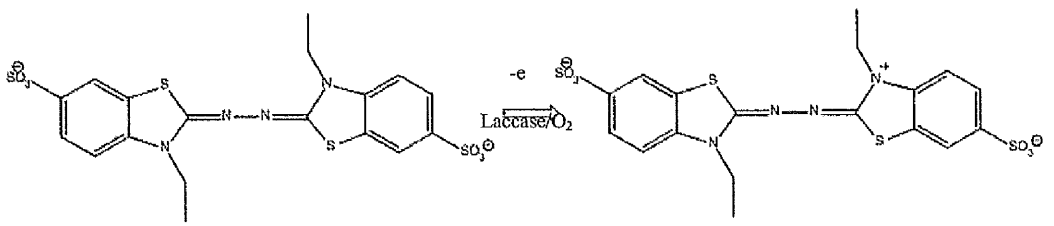


Figure 4.5 Principle of ABTS laccase assay.

In the presence of oxidation catalyst laccase ABTS is becomes oxidised to a metastable radical cation. The appearance of this product can be measured spectrophotometrically at 420 nm.

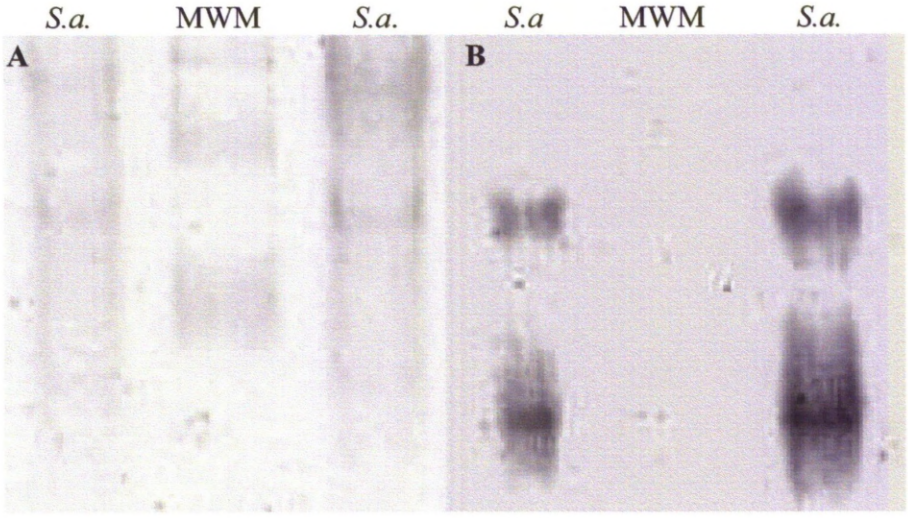


Figure 4.6 Comparison of glycoprotein and *o*-dianisidine activity staining of *S. aeruginosa* extracellular proteins.

Around 10 µg of precipitated protein was then analysed by 9 % native PAGE gel and stained with: glycoprotein strain (A) and activity stain (B): *S.a.* – *S. aeruginosa* protein extract; MWM – Glycoprotein Molecular Weight Marker (negative activity control).

Protein eluted from ConA was then saturated with ammonium sulphate to a final concentration of 0.8 M. All the precipitated proteins were removed by centrifugation and the supernatant was loaded onto hydrophobic interactions column (Phenyl Toyopearl) and eluted stepwise with decreasing ionic strength (Figure 4.9). Various Phenyl Toyopearl fractions were assayed by the ABTS assay, SDS PAGE and native PAGE coupled with *o*-dianisidine staining (Figure 4.7). For sample purity reasons it was decided to focus on two fractions from the HIC: the non-bound fraction and that eluted with 0.064 M $(NH_4)_2SO_4$. Both samples were collected separately, pooled from individual fractions and ultra-filtered using spin concentrators of a molecular weight cut off of 30 kDa in order to remove salts. A non-bound fraction (N) of laccases was buffer exchanged to pH 9.0 during concentration, loaded onto a MonoQ ion exchanger and eluted with increasing ionic strength (Figure 4.10). Protein

from one of the active peaks was found to be over 95% homogenous and used for subsequent analysis. Protein sample (E) eluted from the Toyopearl column was loaded onto Superdex 75 (Figure 4.11) for purification and molecular weight determination, as was the active peak of the N protein eluted from MonoQ at NaCl concentration of 0.4 M. The size exclusion chromatography column was pre-calibrated with molecular weight standards and the estimated molecular weight of both proteins was around 55 kDa. Table 4.1 and Figures 4.12 and 4.13 show selected steps from protein purification experiments and protein purification balance. Alongside the SDS-PAGE analysis native PAGE, coupled with *o*-dianisidine activity staining, was used to asses the purity of isolated proteins. Using this technique it was found that the proteins differ significantly in electrophoretic mobility despite having similar molecular weights, Figure 4.14 summarizes the final steps of purification showing activity staining of purified samples.

Table 4.1. Purification balance of *S. aeruginosa* laccases.

HIC – Phenyl Toyopearl; NB – non bound; EL – eluted; Specific activity is presented in Units per mg of protein i.e μmols of the product formed in one minute by mg of protein.

No.	Purification step	Specific activity of protein [U mg^{-1}]	Enrichment
1	Salting out	9.5	-
2	Q Sepharose	16.16	1.7
3	ConA	155.92	16.4
4	HIC NB	157.71	16.6
5	HIC EL 0.28 M $(NH_4)_2SO_4$	198.00	20.9
6	HIC EL 0.064 M $(NH_4)_2SO_4$	120.13	12.9
7	MonoQ 0.4 M NaCl	323.10	40.0

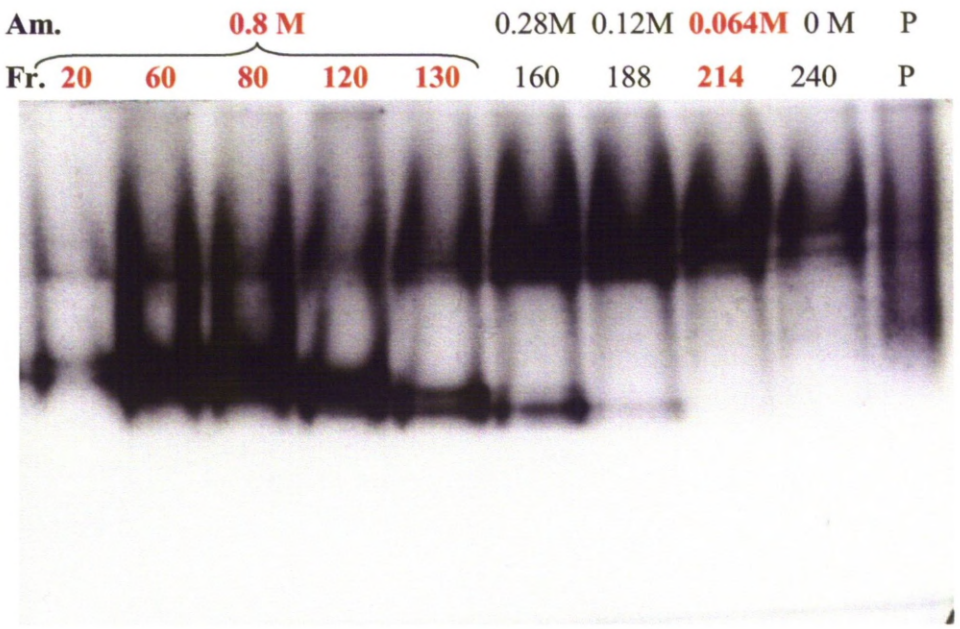


Figure 4.7 Laccase activity elution profile from a Phenyl Toyopearl column.

10 µl of protein was removed from selected fractions and analysed by native PAGE (9 % gel) and *o*-dianisidine staining. The protein samples indicated in red were selected for further purification. Fr. - Fraction number; Am - ammonium sulphate concentration (M); P - protein precipitated by addition of 0.8 M ammonium sulphate

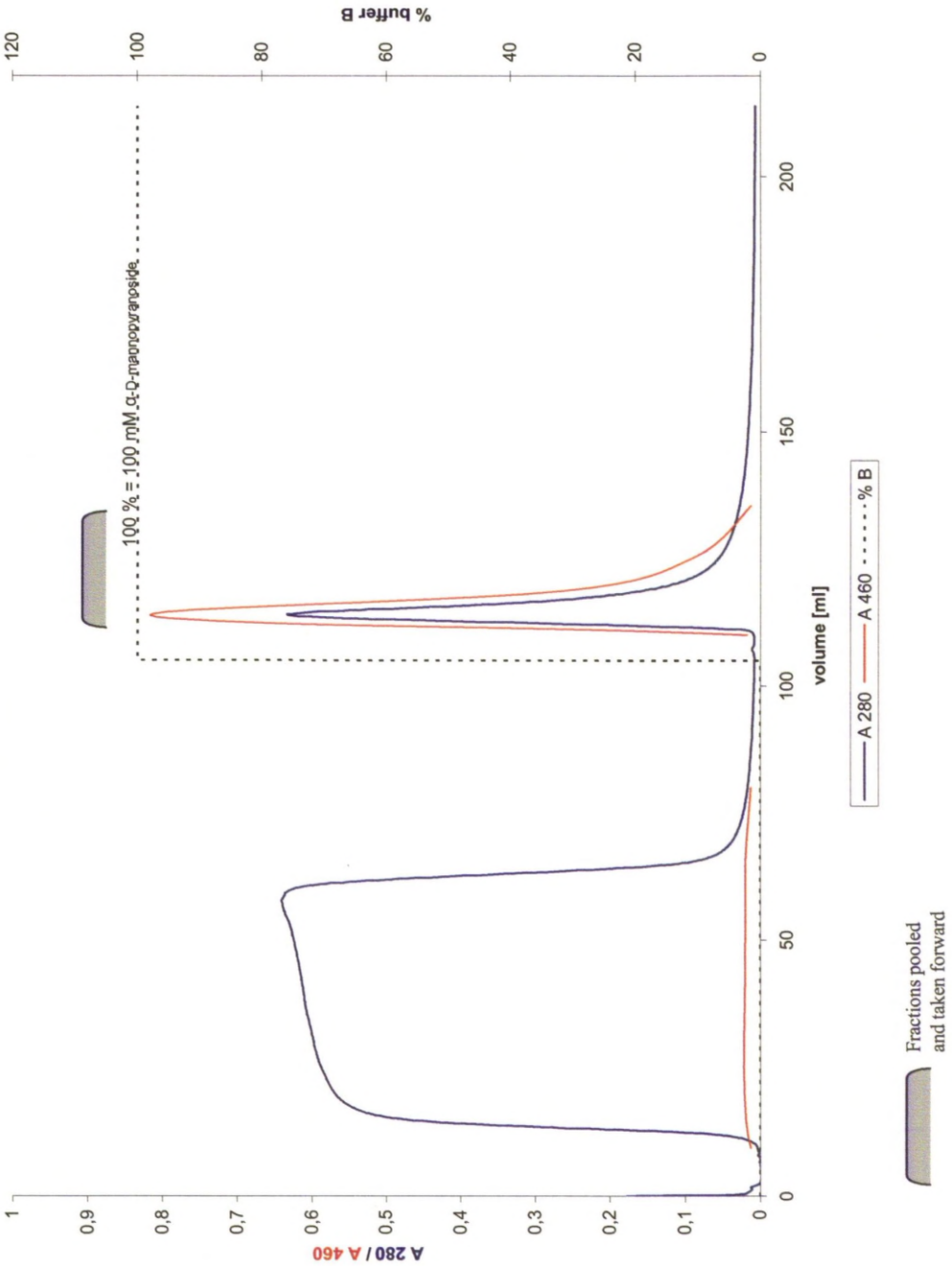


Figure 4.8 ConA agarose chromatography of *S. aeruginosa* laccases.

Depigmented protein sample was loaded onto a preequilibrated ConA column and eluted with 100 mM α-D-methylmannopyranoside in a single step.

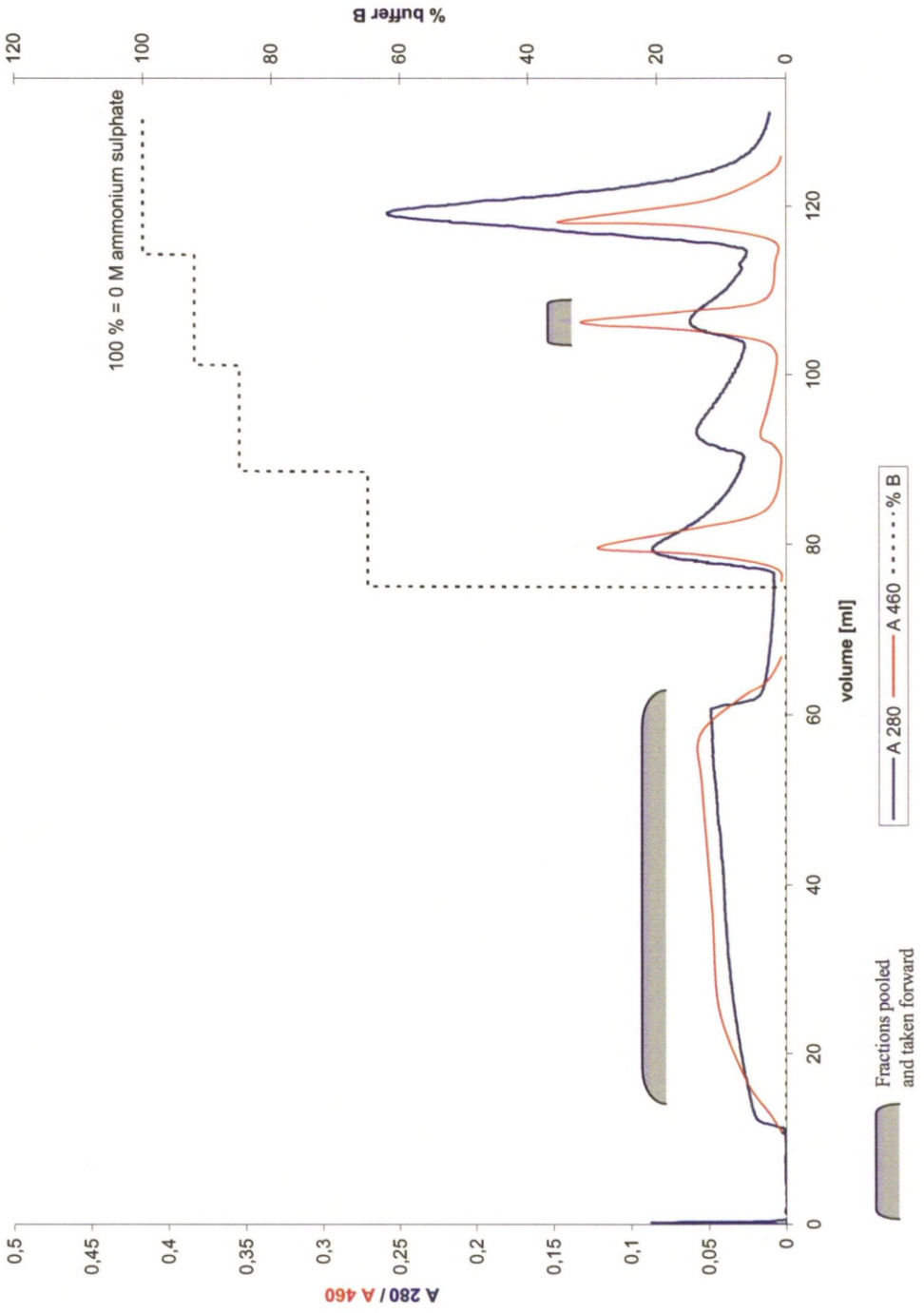


Figure 4.9 HIC chromatography of *S. aeruginosa* laccases.

ConA eluates were loaded onto preequilibrated Phenyl Toyopearl in 0.8 M $(NH_4)_2SO_4$ and eluted step wise with decreasing concentrations of $(NH_4)_2SO_4$.

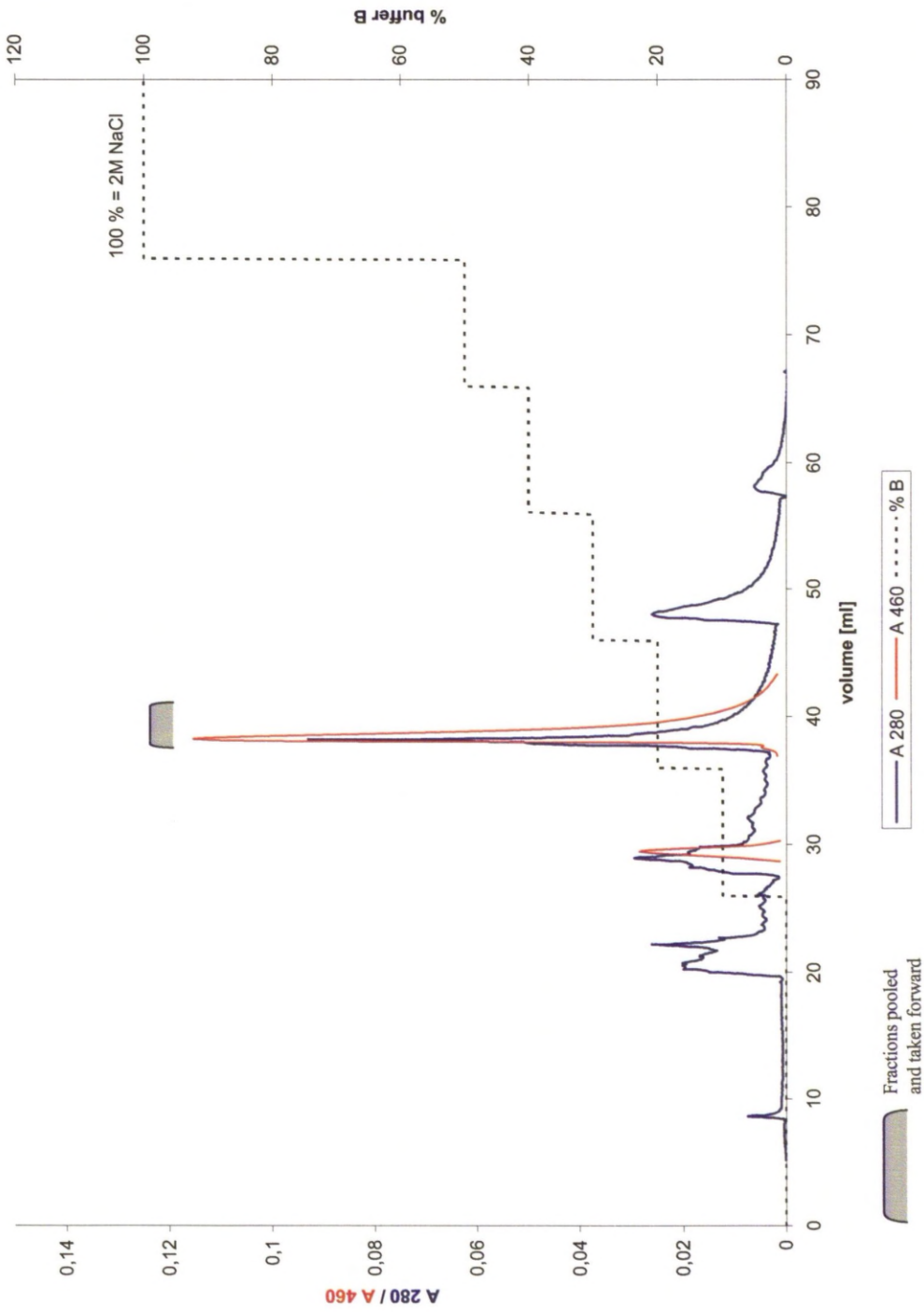


Figure 4.10 Mono Q chromatography of *S. aeruginosa* N laccase.

Non-bound fractions of HIC were loaded onto pre-equilibrated MonoQ column and eluted step-wise with increased concentrations of NaCl.

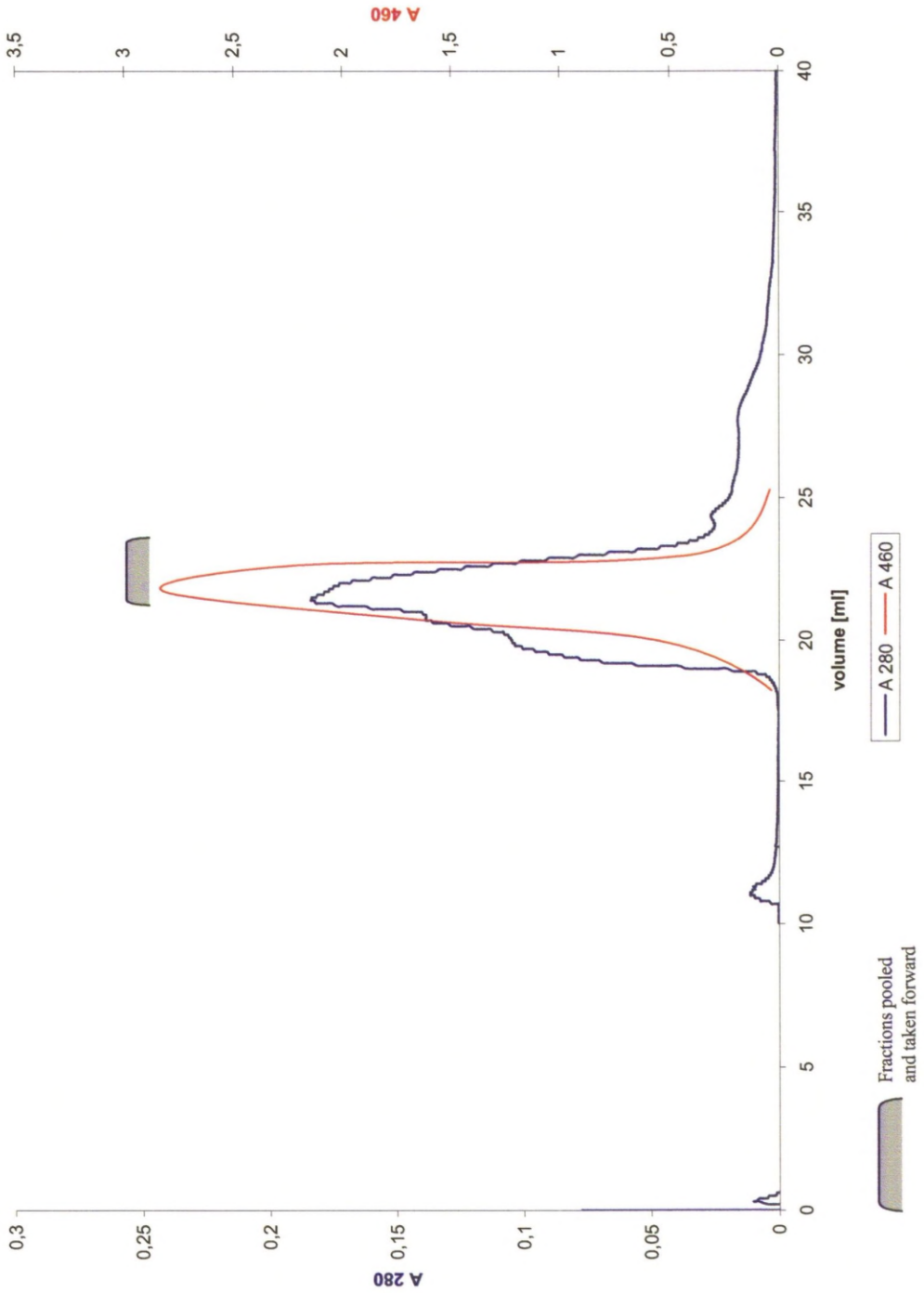


Figure 4.11 Size exclusion chromatography of *S. aeruginosa* E laccase.

Eluted with 0.064 M $(NH_4)_2SO_4$ fractions were loaded onto preequilibrated Superdex 75 column and separated through gel filtration.

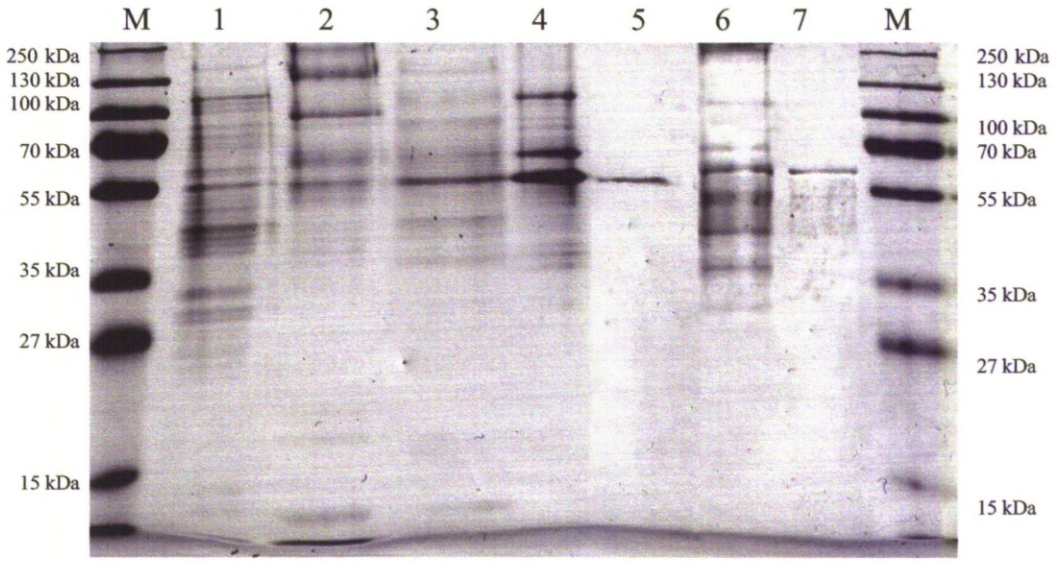


Figure 4.12 Selected steps from the purification of two selected *S. aeruginosa* laccases (SDS-PAGE).

M – molecular weight marker; 1 – Extract; 2 – Q sepharose eluted fractions; 3 – ConA eluted fractions (Figure 4.8); 4 – HIC non bound fractions (Figure 4.9); 5 – 0.4 M NaCl MonoQ eluted fractions (Figure 4.10); 6 – 0.28 M $(NH_4)_2SO_4$ HIC eluted fractions (Figure 4.9); 7 – 0.064 M $(NH_4)_2SO_4$ HIC eluted fractions (Figure 4.9). Separation was performed on 12 % gel.

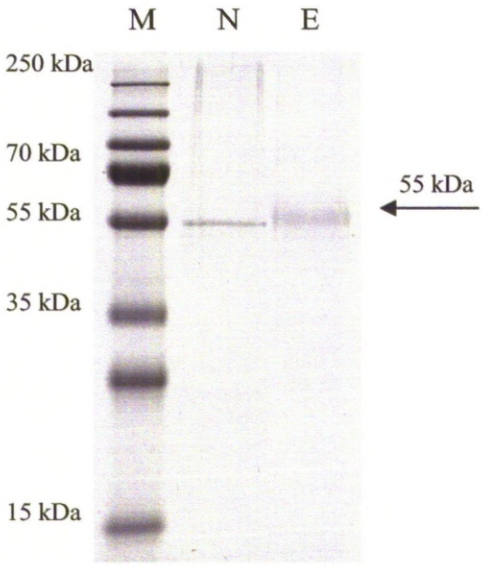


Figure 4.13 Purification of two selected *S. aeruginosa* laccases (SDS-PAGE).

M – molecular weight marker plus; N – N laccase after last step of purification; E – E laccase after last step of purification Separation was performed on 12 % gel.

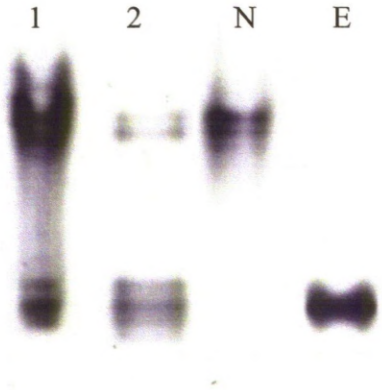


Figure 4.14 Purification of two selected *S. aeruginosa* laccases (native PAGE *o*-dianisidine activity staining).

1 – laccases after ConA purification step (Figure 4.8); 2 – HIC non bound fractions (Figure 4.9); E – E laccase after last step of purification; N – N laccase after last step of purification. Separation was performed on 9 % gel.

4. Basic biochemical properties of *S. aeruginosa* enzymes.

The purified protein samples N (non-bound to HIC) and E (eluted from HIC) were used to assess the basic biochemical properties of *S. aeruginosa* laccases. The effect of both temperature and pH on protein stability and enzymatic activity was tested with ABTS. Both enzymes exhibit moderate optimal temperatures of activity (Figure 4.15) and acidic optimal pH of activity (Figure 4.17). Both enzymes are 100 % stable for 1 hour at 37 °C and retain over 60 % of activity after 1 hour incubation at 60 °C (Figure 4.16). Both enzymes seem to be more stable in alkaline pH than in acidic conditions, the N enzyme exhibits very high stability even after prolonged time incubation (1 week) at pH 9 (Figure 4.18). Neither of the proteins showed any tyrosinase activity with L-DOPA and L-tyrosine, this feature is important for distinguishing real laccases from tyrosinases that show very similar substrate specificity with phenolic substrates (Flurkey et al., 1995). Another finding that confirms that both enzymes are laccases is the fact that both proteins perform much worse in degassed buffers than aerated, as this confirms the enzymes are dependent on molecular oxygen as an electron acceptor, a feature characteristic of laccases.

The analysis of biochemical properties of these laccases was expanded by Catherine Ann Houghton in her undergraduate project titled: "Biochemical characterisation of laccases from *S. aeruginosa*". The project involved more detailed investigation of substrate specificity, the effect of inhibitors and possible applications of laccases in the textile industry through decolourisation of anthraquinone and azo dyes.

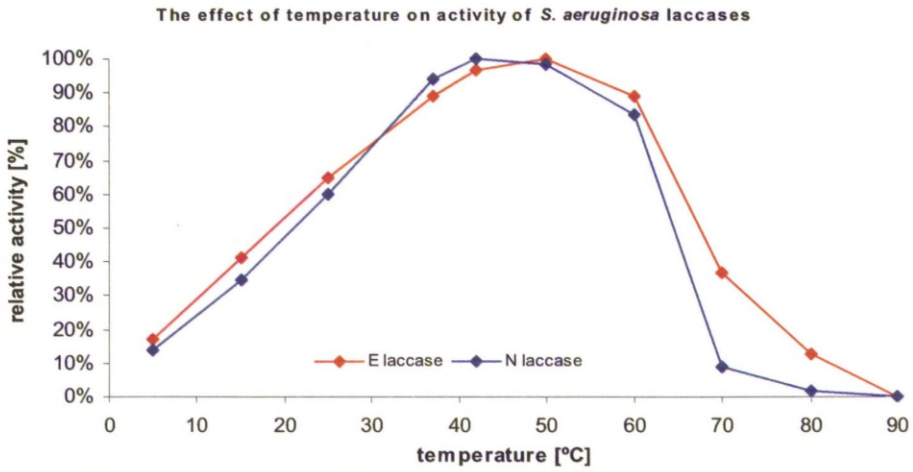


Figure 4.15 Effect of temperature on activity of *S. aeruginosa* laccases.

Enzymatic activity of laccases at various temperatures was measured in triplicatesby incubating the sample for the time of the assay (15 min). Relative activity as a percentage of maximal activity is plotted against assay temperature.

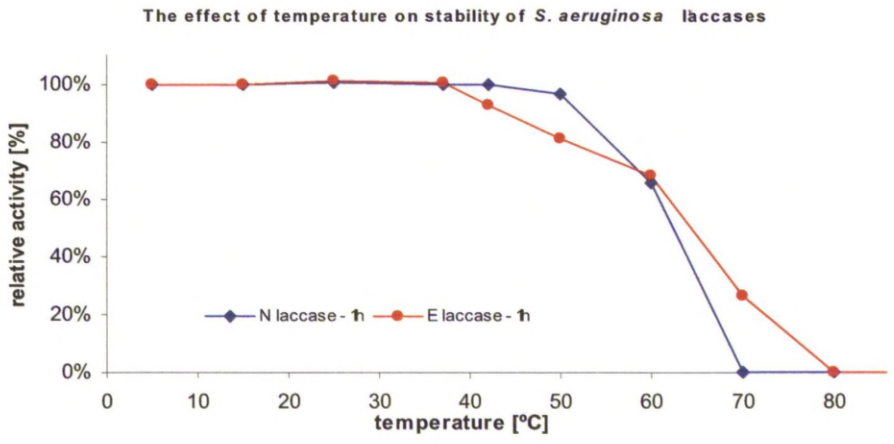


Figure 4.16 Effect of temperature on stability of *S. aeruginosa* laccases.

Thermostability of laccases was measured in triplicates by incubating the sample at different temperatures for 1 hour before performing an assay at standard conditions (42 °C, 15 min). Relative activity as a percentage of maximal activity is plotted against assay temperature

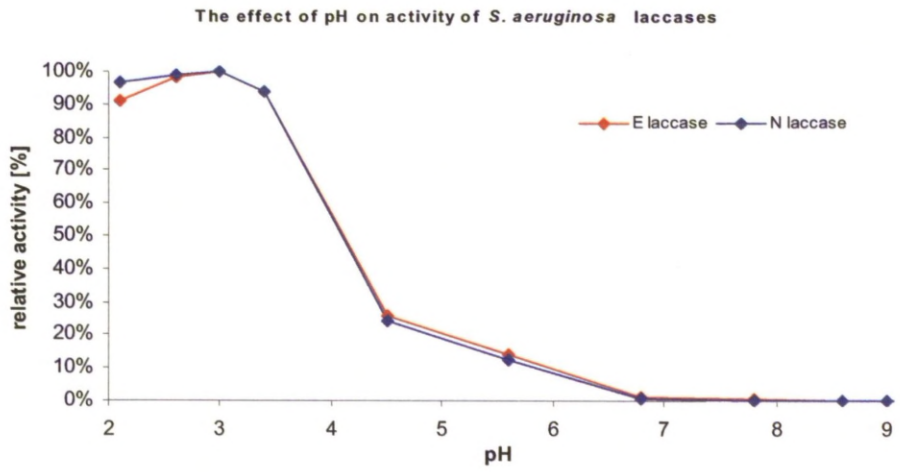


Figure 4.17. Effect of pH on activity of *S. aeruginosa* laccases.

Enzymatic activity of laccases at various pH was measured in triplicates by incubating the sample for the time of the assay (15 min). Relative activity as a percentage of maximal activity is plotted against pH of the assay.

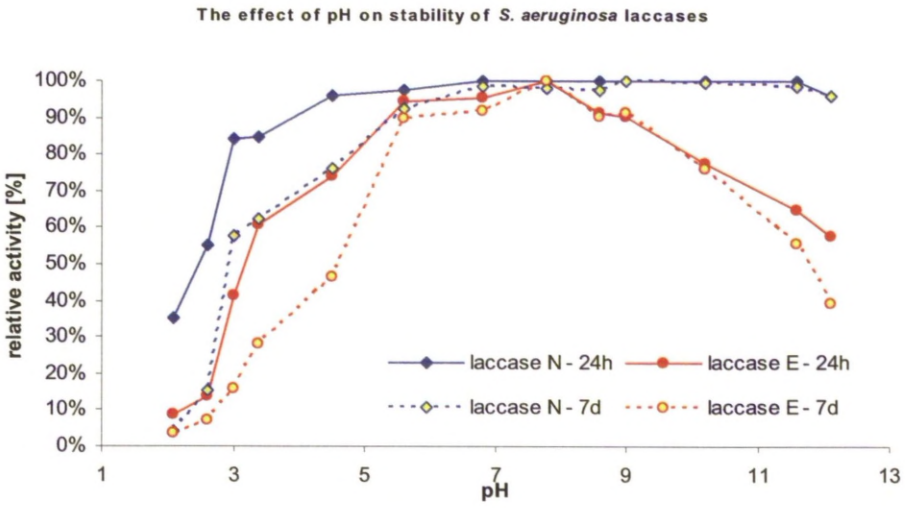


Figure 4.18 Effect of pH on stability of *S. aeruginosa* laccases.

Stability of laccases at various pH was measured in triplicates by incubating the sample at different pH for 24 hours and one week before performing an assay at standard conditions (42 °C, 15 min). Relative activity as a percentage of maximal activity is plotted against pH of the assay.

Table 4.2 The summary of basic biochemical properties of *S. aeruginosa* laccases.

Feature [ABTS test]	**N laccase**	**E laccase**
Optimal temperature of activity	42 °C	50 °C
Thermostability	100 % stable for 1 hour at 50 °C Inactivated after 1 hour incubation at 70 °C	100 % stable for 1 hour at 37 °C Inactivated after 1 hour incubation at 80 °C
Optimal pH of activity	3.0	3.0
pH stability	> 90% stable for a week at 5.5 to 12 at room temperature	> 90% stable for a week at 5.5 to 9 at room temperature

5. Spectroscopic properties of *S. aeruginosa* enzymes.

The UV/VIS spectrum of the purified proteins was assessed to determine whether they showed any typical characteristics of some groups of oxidases i.e. the presence of catalytic copper. The spectrum of purified proteins is shown below in Figure 4.19.

Despite showing the substrate specificity of laccases, along with the appropriate size and behaviour the isolated proteins did not seem to show any spectroscopic characteristics of typical blue laccases i.e. the characteristic band of T1 copper at 600 nm (Baldrian, 2006). After an extensive literature search it was decided to annotate them as members of a rare group of oxidoreductases called yellow laccases.

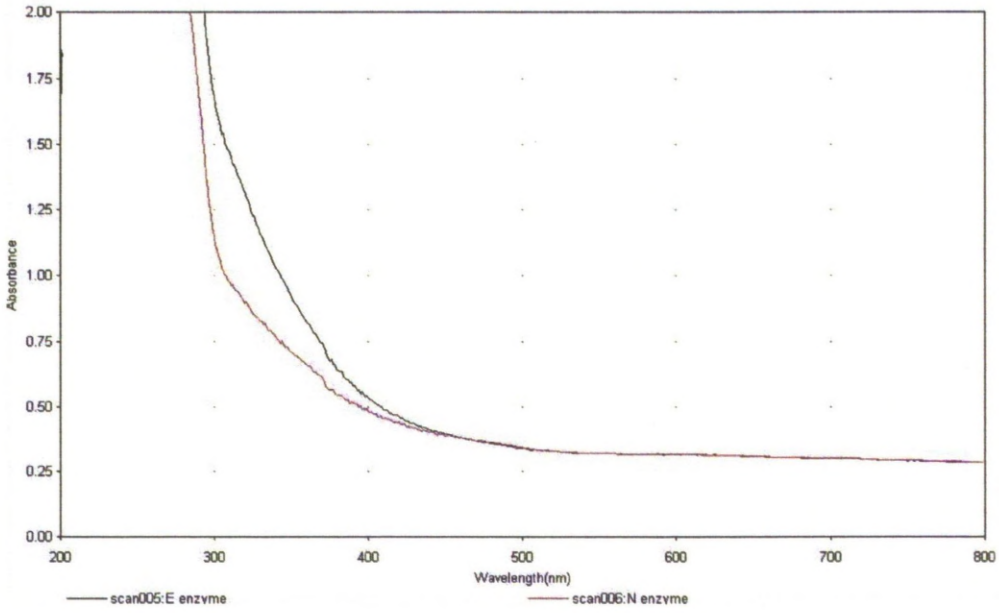


Figure 4.19 UV-VIS spectrum of purified laccases N (red) and E (blue).

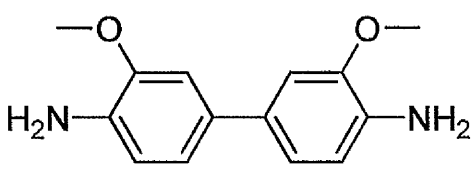
The UV-VIS spectra of the purified samples of yellow laccases from *S. aeruginosa,* showing the lack of the characteristic absorption band at 600 nm which is responsible for the blue colour of blue laccases.

6. Advanced biochemical properties of *S. aeruginosa* enzymes (partially adapted from C.A. Houghton's project report)

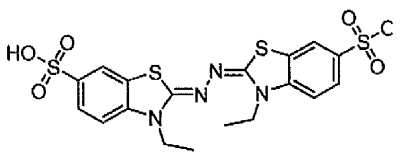
The Catherine Ann Houghton honours project included more detailed investigation of substrate specificity, kinetics and inhibitory effect of many small molecular weight compounds. Attempts to find an industrial application in the textile industry were also made.

6.1. Detailed substrate specificity.

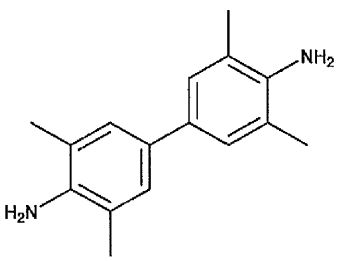
A number of popular laccase substrates were tested and selected for analysis on the basis of the literature (Morozova et al., 2007, Giardina et al., 2010, Baldrian, 2006). The chemical formulae of these substrates are presented on Figures 4.20 and 4.21 below.



o-Dianisidine



ABTS



Tetramethylbenzidine

Figure 4.20 Non-phenolic substrates used in activity assay with *S. aeruginosa* laccases.

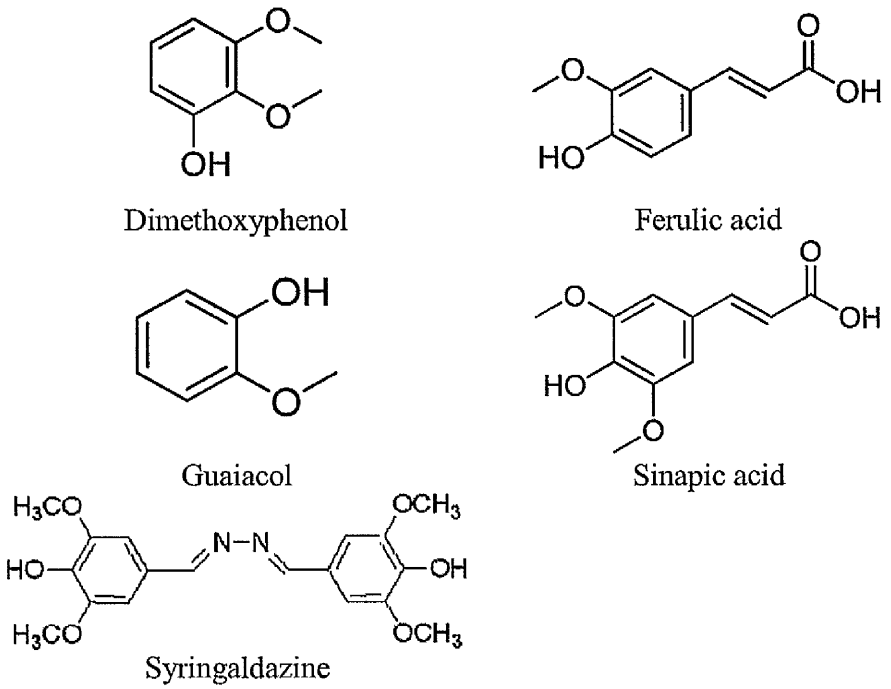


Figure 4.21 Phenolic substrates used in activity assay with laccases.

The substrates were tested at pH ranging from 2 to 11.6, the substrates included: synthetic non phenolic substrates: ABTS, *o*-dianisidine, tetramethylbenzidine (TMB), syringaldazine; non lignin phenolic derivatives: dimethoxy phenol (DMP), guaiacol and lignin-like phenolic compounds: ferulic acid and sinapic acid. Figures 4.22 and 4.23 present the pH activity profiles of those substrates for N laccase and E laccase respectively.

It was shown that both enzymes exhibit similar substrate specificities with non phenolic ABTS being the substrate showing the highest specific activity and two lignin-like phenolic compounds sinapic and ferulic acids showing very high activities as well, however syringaldazine does not seem to get readily oxidised by the laccases of *S. aeruginosa*. It is worth mentioning that each of these above mentioned groups of compounds show similar optimal pH of activity i.e. synthetic substrates like ABTS, TMB and *o*-dianisidine are best oxidised at low pH. The activity drops

significantly above a pH of 4.5, it should also be mentioned that they exhibit a rather narrower window of activity than most of the phenolic derivatives. Lignin derived phenolics, including syringaldazine, posseses a broader pH of activity and are oxidised in a pH more similar to that of the natural environment. The specific activities in optimal pH differ between various substrates even a hundredfold with methoxy substituted phenols (DMP and guaiacol) being the least specific substrates for the tested laccases. Once the pH activity profiles of both laccases were established, and optimal pH for each substrate was selected, it was attempted to compare the specific activities of enzymes measured for each substrate at its optimal pH of activity (Figure 4.24) and their effective pH range (Table 4.3).

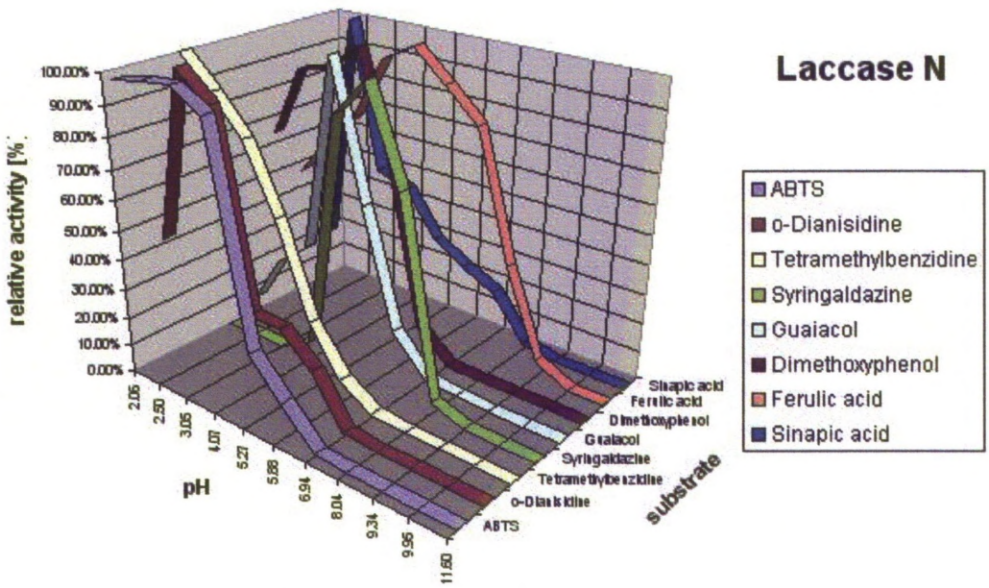


Figure 4.22 The pH activity profiles of the N laccase towards various substrates at 42 °C.

The figure shows that N laccase of *S. aeruginosa* performs best at low pH with non-phenolic substrates like: ABTS, tetramethylbenzidine and *o*-dianisidine, at higher pH activity significantly decreases. Towards phenolic substrates (dimethoxyphenol, ferulic acid), the laccase exhibits higher optimal pH of activity and the enzyme shows more tolerance to non-optimal pH values.

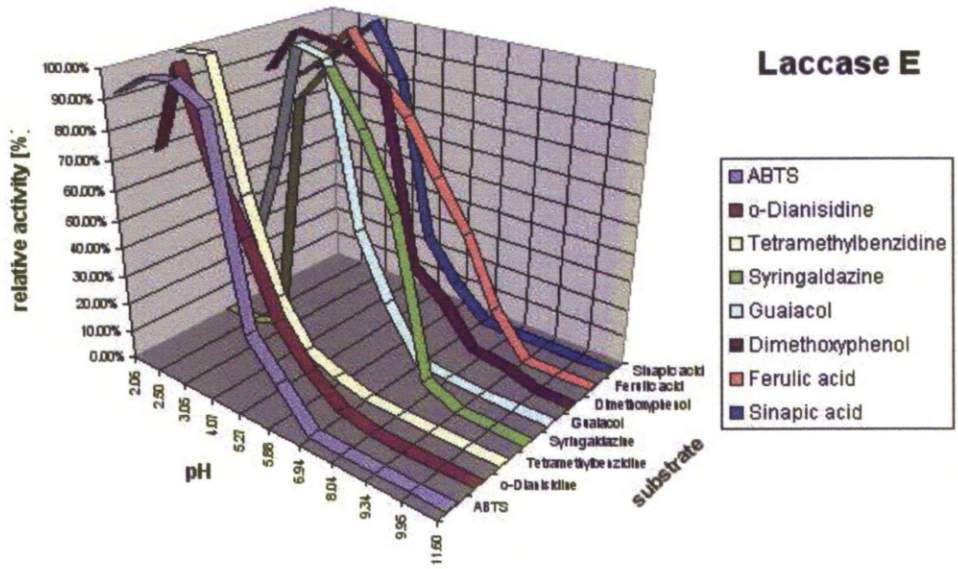


Figure 4.23 The pH activity profiles of E laccase towards various substrates at 42 °C.

The figure shows that E laccase of *S. aeruginosa* performs best at low pH with non-phenolic substrates like: ABTS, tetramethylbenzidine and *o*-dianisidine, at higher pH activity significantly decreases. Towards phenolic substrates (dimethoxyphenol, ferulic acid), the laccase exhibits higher optimal pH of activity and the enzyme shows more tolerance to non-optimal pH values.

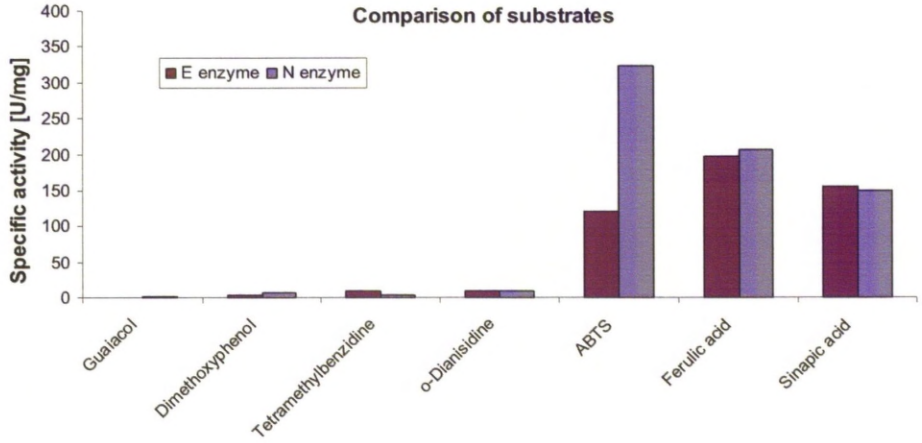


Figure 4.24 Summary of the N and E laccases specific activities towards selected substrates at their optimal pH at 42 °C showing that ABTS, ferulic and sinapic acids are the best substrates for both laccases.

Table 4.3 Summary of the N and E laccase optimum conditions and activities.

eff. pH – the effective pH range – range of pH where enzyme exhibits at least 70 % of its maximal activity
SA – specific activity of pure protein [$U \cdot mg^{-1}$]

	Laccase N		
Substrate	**eff. pH**	**Opt. pH**	**SA**
Guaiacol	3.5 – 5.0	4.07	2.35
DMP	2.2 – 5.2	4.07	8.16
TMB	2.0 – 3.4	2.06	3.69
o-DA	2.2 – 3.4	2.50	8.69
ABTS	2.0 – 4.5	3.05	323.10
Syringaldazine	5.0 – 6.9	5.88	8.51
Ferulic acid	3.0 – 7.2	5.27	206.89
Sinapic acid	2.2 – 2.8	2.50	150.55
	Laccase E		
Substrate	**eff. pH**	**opt. pH**	**SA**
Guaiacol	2.6 – 4.7	3.05	0.83
DMP	2.0 – 5.5	2.50	3.39
TMB	2.0 – 2.8	2.50	8.34
o-DA	2.0 – 3.0	2.50	8.50
ABTS	2.0 – 4.5	3.05	120.13
Syringaldazine	3.8 – 6.6	5.27	8.55
Ferulic acid	2.0 – 5.4	3.05	196.97
Sinapic acid	2.0 – 4.4	3.05	155.56

6.2. Kinetic parameters of *S. aeruginosa* laccases.

Once the optimal conditions for each substrate and enzyme were identified, kinetic constants for a few selected substrates: ABTS, *o*-dianisidine and syringaldazine, were investigated. A summary of this kinetic data is presented in Table 4.4.

Table 4.4 Summary of kinetic data for selected substrates and at 42 °C and given pH.

	Laccase N				
Substrate	**pH**	**V_{max} [µmol min $^{-1}$]**	**K_m [mM]**	**k_{cat} [s^{-1}]**	**k_{cat}/K_m**
o-DA	2.5	117.4	10.99	14,405	1,311
ABTS	3.0	285.3	180	35,016	194,533
Syringaldazine	4.0	25.4	0.72	3,122	4,336
	Laccase E				
Substrate	**pH**	**V_{max} [µmol min $^{-1}$]**	**K_m [mM]**	**k_{cat} [s^{-1}]**	**k_{cat}/K_m**
o-DA	2.5	169.3	4.52	7,643	1,691
ABTS	3.0	254.9	62	11,508	185,613
Syringaldazine	4.0	18.5	0.48	833	1,735

Kinetic analysis shows that ABTS is the most specific substrate for both laccases of *S. aeruginosa*. The specificity constant k_{cat}/K_m is 5-10 fold greater for ABTS then for either *o*-dianisidine or syringaldazine.

6.3. The effect of inhibitors on enzymatic activity of *S. aeruginosa* laccases.

Once the kinetic parameters were established it was attempted to test the inhibitory effects of small compounds on the activity of *S. aeruginosa*

laccases. A number of popular laccase inhibitors were tested and selected for analysis on the basis of literature (Morozova et al., 2007, Giardina et al., 2010, Baldrian, 2006) to confirm that the purified proteins are indeed laccases. The effect of various inhibitors during a 1 hour long incubation was tested using ABTS as a substrate. The preliminary data obtained, on partially purified proteins, showed inhibition with: thiols, halides, hydroxylamine and sodium azide (results not shown) therefore during C.A. Houghton's project it was attempted to obtain more detailed information about the inhibition. Table 4.5 summarizes these findings showing IC_{50} values for each tested inhibitor (i.e. the concentration of inhibitor that inactivates 50 % of the enzyme during 1 hour of incubation).

The result shows that the enzymes of *S. aeruginosa* exhibit typical laccase characteristics during inhibitory tests, in that they are significantly inhibited by thiols, sodium azide, *L*-ascorbic acid and *L*-cysteine and to a lesser extend hydroxylamine, halides and EDTA (Baldrian, 2006).

Table 4.5 Summary of kinetic data for selected substrates at following conditions: incubation buffer pH 6.5, reaction buffer pH 3.0, 42 °C, 1 h.

	Laccase N	**Laccase E**
Inhibitor	**IC_{50} [mM]**	**IC_{50} [mM]**
DTT	1.00	7.00
βmerEtOH	0.50	0.50
Thioglycolic Acid	1.00	10.00
Sodium Azide	0.30	0.30
Hydroxylamine	30.00	150.00
L-ascorbic Acid	0.05	0.40
L-Cysteine	4.00	1.00
EDTA	160.00	175.00

7. Decolourisation of textile dyes as a potential application of *S. aeruginosa* laccases. (adapted from C.A. Houghton's project report)

There are over 100,000 different dyes used by various industries (Chairattanamanokorn et al., 2006). Many of these compounds are resistant to degradation and are discharged to sewage systems on a daily basis. Efficient degradation of these compounds presents a serious challenge. It is generally accepted that laccases are one of the enzymes that can tackle this problem (Lopez et al., 2002, Hilden et al., 2009). In C.A. Houghton's project both purified laccases were tested in decolourisation of selected azo and anthraquinone dyes that represent two of the major groups of dyes used in the textile industry. The dyes for decolourization experiments were selected on the basis of the literature review (Giardina et al., 2010, Miele et al., 2009, Boonyakamol et al., 2009). Each of the experiments was performed over the period of 24 hours and assessed in hourly intervals using 0.08 ABTS unit of enzyme at pH 3.0, 42 °C. The time course results of this experiment are presented in Figure 4.27 and 4.28 for N and E laccases respectively, whereas Figure 4.29 shows actual colour changes observed after 3 and 24 hours of decolourisation. Figures 4.25 and 4.26 present chemical structures of dyes used in this experiment.

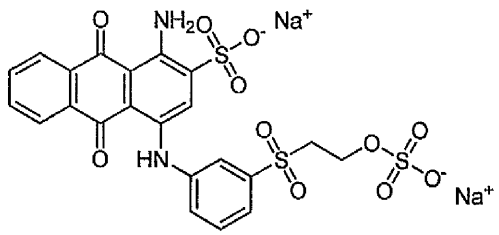


Figure 4.25 Anthraquinone dye Remazol Brilliant Blue R used in decolourisation experiment.with *S. aeruginosa* laccases.

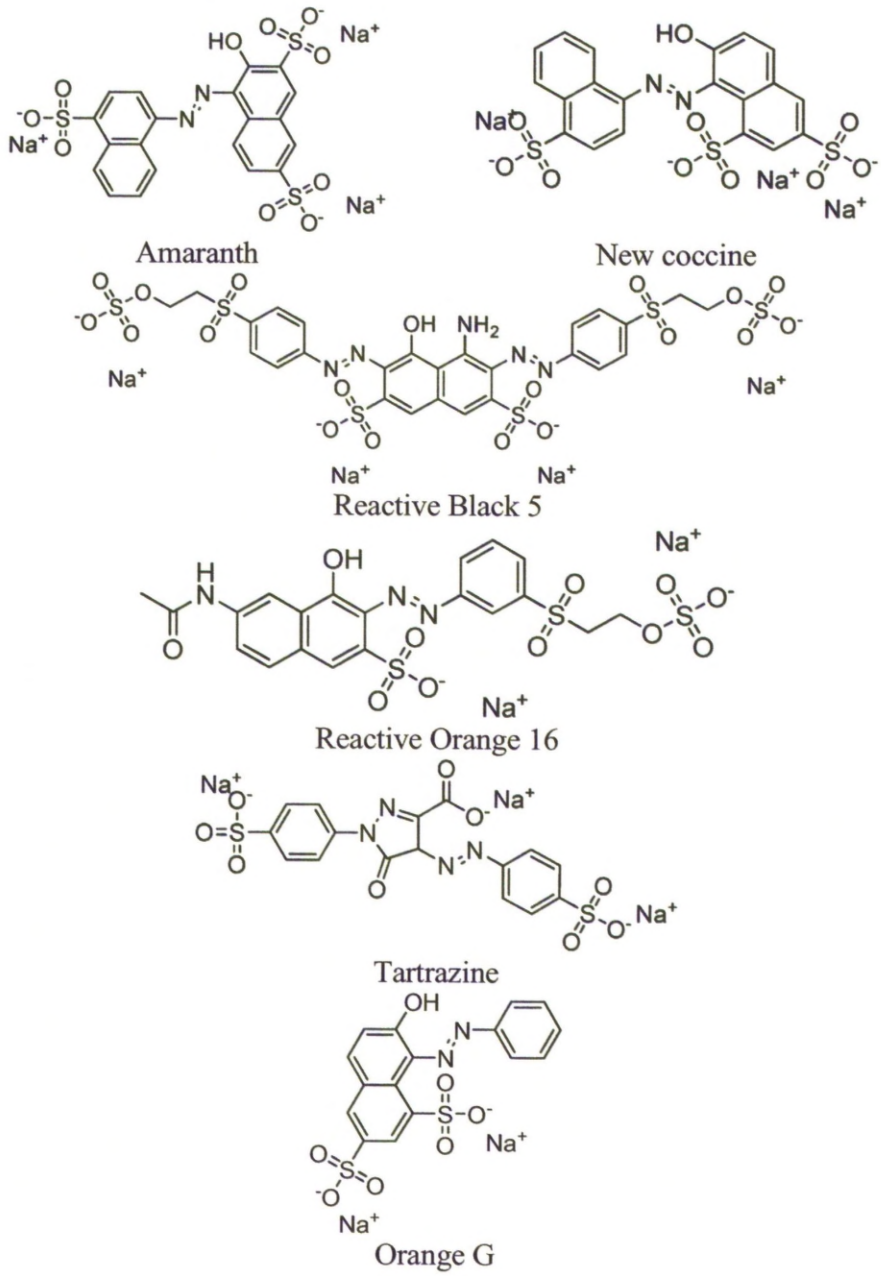


Figure 4.26 Azo dyes used in decolourisation experiment with *S. aeruginosa* laccases.

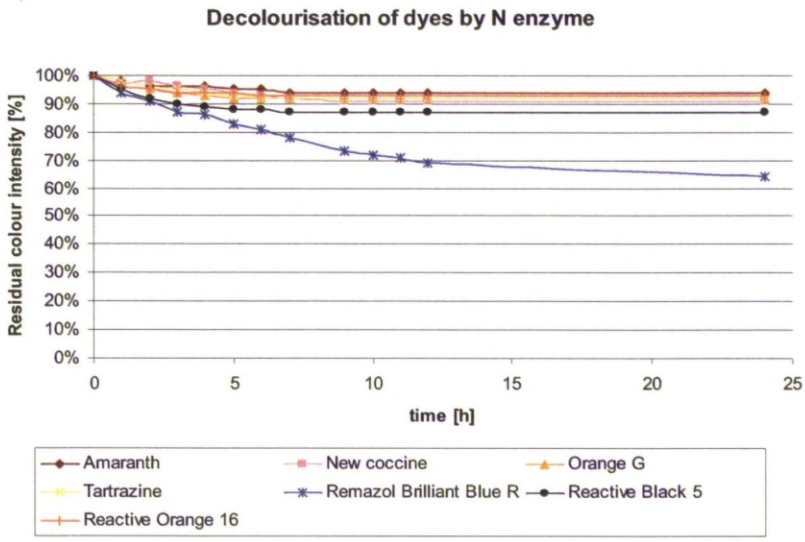


Figure 4.27 Time course of textile dye decolourisation with the N laccase.

Decolourisation of various azo and anthraquinone dyes was performed with laccase N. Results of decolourisation show that laccase N of *S. aeruginosa* is a weak dye decolouriser. Decline of decolourisation rate after 5 hours is caused by enzyme deactivation due to low pH of the assay (pH 3).

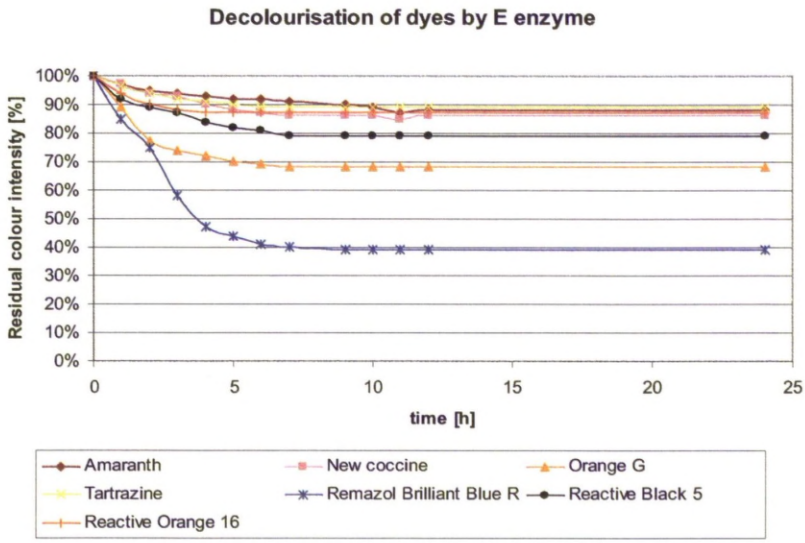


Figure 4.28 Time course of textile dye decolourisation with the E laccase.

Decolourisation of various azo and anthraquinone dyes was performed with laccase E. Results of decolourisation show that laccase E of *S. aeruginosa* is good decolouriser of Remazol Brilliant Blue R. Decline of decolourisation rate after 5 hours is caused by enzyme deactivation due to low pH of the assay (pH 3).

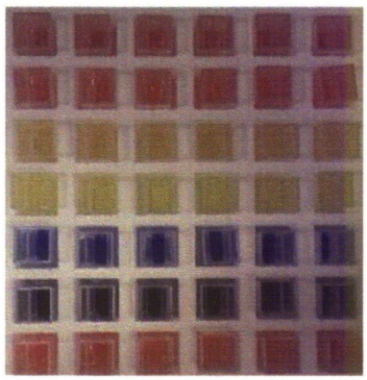
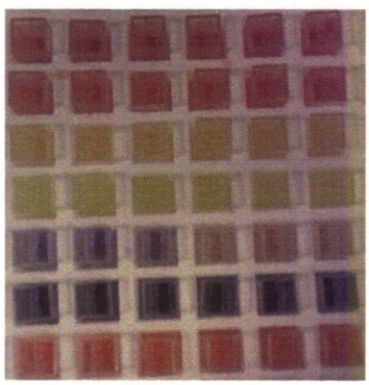
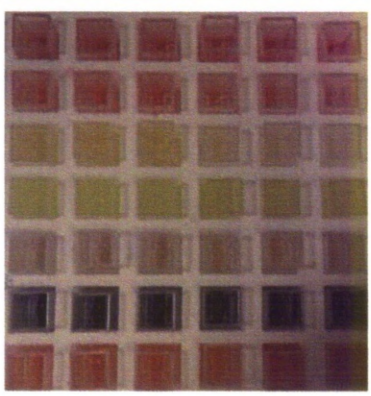
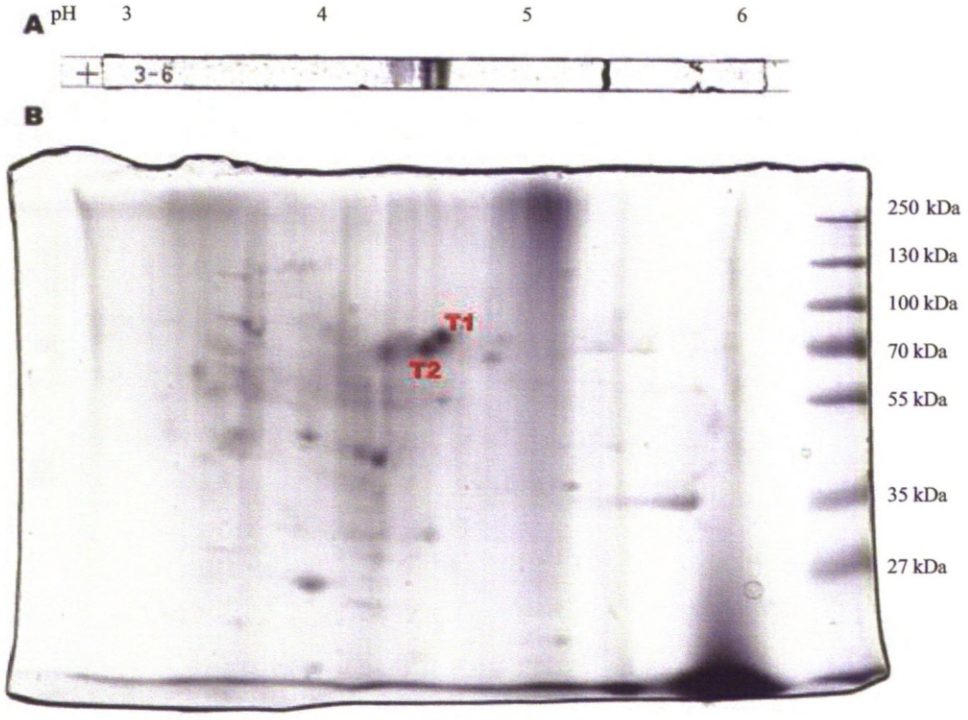
Time	Dye	N laccase	E laccase
0 h	Amaranth		
	New Coccine		
	Orange G		
	Tartrazine		
	Remazol Brilliant Blue R		
	Reactive Black 5		
	Reactive Orange 16		
3 h	Amaranth		
	New Coccine		
	Orange G		
	Tartrazine		
	Remazol Brilliant Blue R		
	Reactive Black 5		
	Reactive Orange 16		
24 h	Amaranth		
	New Coccine		
	Orange G		
	Tartrazine		
	Remazol Brilliant Blue R		
	Reactive Black 5		
	Reactive Orange 16		

Figure 4.29 Comparative enzymatic decolourisation of textile dyes with both *S. aeruginosa* laccases after 0, 3 and 24 hours of decolourisation.

8. Proteomic analysis of *S. aeruginosa* laccases.

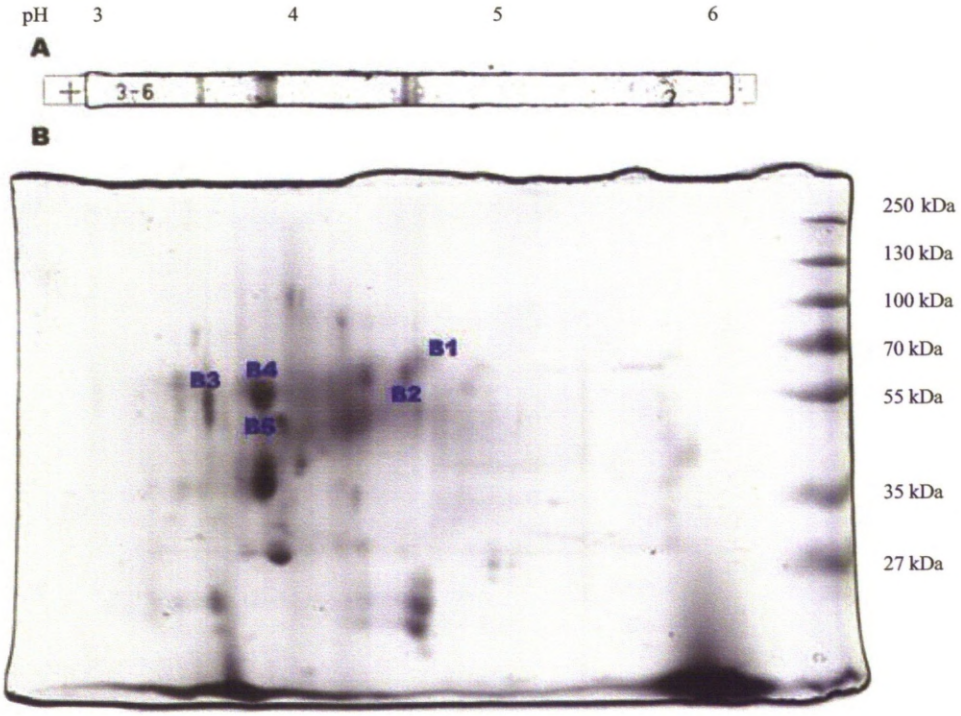
A proteomic analysis was performed in order to further characterise the two laccases and obtain partial amino acid sequences. To achieve this, the ConA flow-through and eluates were subjected to 2D analysis. The concentrated protein sample was run in native IEF strips in duplicate; one of the strips was stained with *o*-dianisidine for laccase activity and the other run in a second dimension in the presence of SDS. After a few initial trials the pH range chosen for the first dimension separation was set in the acidic range 3-6 as this gave the best resolution in the isoelectric range among all the tested IEF strips. The resulting 2D gels are presented in Figures 4.30 and 4.31 for non bound and eluted fractions respectively. Seven spots corresponding to the band of activity, and showing the expected molecular weight of around 60 kDa, were excised and digested in gel with trypsin and V8 protease. The flow through samples (denoted T) showed two major spots on the 2D gel, the samples that were bound to the ConA medium and subsequently eluted (denoted B) showed 5 spots that fulfil the criteria. The resulting peptides were ran on a tandem mass spectrometer by Mr Mark Prescott and sequenced manually. The map of peptides corresponding to particular spots on the gel is presented in below Figures 4.30 and 4.31.



Spot	Peptide	Abundance (number of sequences)
T1	LTVKVPLRT	9
	TEYNVNV	2
	SSPATL	1
T2	LTVKVPLRT	9
	TSPLSVLTVL	1

Figure 4.30 Semi native 2D gel representing the protein separation of concentrated ConA column flow through.

ConA flow through was concentrated and loaded onto IEF strip pH 3-6. After focusing, the sample was run on 10 % SDS gel in the second dimension to separate proteins according to their molecular weight. A - *o*-dianisidine activity staining of replica IEF strip; B-2D gel stained with Gel Code Blue Protein Stain. The table below the gel presents peptide sequences obtained from mass spectrometric sequencing of spots indicated on the gel.



Spot	Peptide	Abundance (number of seqences)
B1	LTVKVPLRT	9
	VVNATAV	5
B2	GPTTNLYL	1
B4	GPNANLFLQNK (GPSTNLFIENK)	2
	KQVWNVIV	1
B5	GPNANLFLQNK (GPSTNLFIENK)	2

Figure 4.31 Semi native 2D gel representing the protein separation of concentrated ConA column eluate.

ConA eluate was concentrated and loaded onto IEF strip pH 3-6 after focusing sample was run on 10 % SDS gel in the second dimension to separate proteins according to their molecular weight. A - *o*-dianisidine activity staining of replica IEF strip; B-2D gel stained with Gel Code Blue Protein Stain. The table below the gel presents peptide sequences obtained from mass spectrometric sequencing of spots indicated on the gel.

When comparing Figures 4.30 and 4.31 one can observe that peptide LTVKVPLRT is common for spots T1, T2 and B1. Another peptide that can be found in more than one protein is GPNANLFLQNK (or GPSTNLFIENK the spectrum was inconclusive, and both sequences gave an equally probable result) that is present in B4 and B5.

Sequence similarity searches were performed on all peptide sequences with PAM 40 (short sequence specific) matrix of BLASTP program to find any homology to laccases in databases. Unfortunately the peptides did not show any sequence similiarity to known laccases.

Two other observations were made; the first being that, very similar peptides were found from different spots suggesting that those sequences were different isoforms of the same protein or that the proteins very closely related to one another. Similar observations were made by Lisova et al. regarding proteomic analysis of *Cerena unicolor* (Lisova et al., 2010). The second observation appears to show that the isoforms differ in their degree of glycosylation since similar spots on 2D gels gave identical peptide matches from both the flow through and eluate of ConA column. The peptide that confirms this hypothesis is LTVKVPLRT. Despite having only very short stretches of protein sequence the attempts were made to use this information to design primers in order to obtain the coding sequences from cDNA library of *S. aeruginosa*. These attempts are presented in Chapter 5.

9. Summary.

As a continuation of Moore's work attempts were made to isolate and characterise a putative haloperoxidase of *S. aeruginosa*, however a combination of technical difficulties rendered this task impossible and even raised the question of whether such enzyme ever existed in this organism. As a result the efforts focused on isolating laccases of *S. aeruginosa* and analysis of their iso and glycoforms. Two laccases were purified using

chromatographic methods and analysed biochemically. Enzymes show typical substrate specificities of laccases with maximal substrate specificity towards synthetic substrate ABTS and the lignin building blocks: ferulic and sinapic acids. The character of laccases was also confirmed by inhibitory tests that showed inhibition patterns with hydroxylamine, azides and halides (typical for other laccases). The analysis of spectroscopic properties suggested that both proteins belong to the family of yellow laccases due to the lack of a characteristic absorbance peak at 600 nm. They are a rare type of enzymes that are not yet fully understood, but some theories suggest that they have a mediator molecule permanently bound to T1 copper site (Leontievsky et al., 1997b). Proteomic analysis of various laccase isoforms suggests that the array of laccases secreted by *S. aeruginosa* is larger than two and the presence of various protein iso and glycoforms is expected. It was also found that one of those laccases (laccase E) is an efficient degrader of anthraquinone dye Remazol Brilliant Blue R. The unusual spectroscopic properties of *S. aeruginosa* laccases and their isoform abundance raised questions about genetic similarity of the genes. The next step was an attempted isolation of laccase gene(s) through *S. aeruginosa* cDNA library screening.

Chapter 5

THE PURSUIT OF THE GENES OF *S. AERUGINOSA* LACCASES USING CONVENTIONAL METHODS.

1. Introduction.

Before the characterisation *of S. aeruginosa* laccases was completed, the identification of the genes encoding those proteins was attempted. Many different methodologies were tried, divided into two main groups: activity-based screens of *S. aeruginosa* cDNA libraries in different organisms, and protein sequence-based approach that involved the design of degenerate oligonucleotide primers/probes based on amino acid sequences obtained from mass spectrometry analysis. The former approach involved colorimetric plate screens, the latter experiments were based on PCR and hybridisation of nucleic acids.

2. Creation of cDNA libraries of *S. aeruginosa.*

2.1. Preparation of unidirectional cDNA from *S. aeruginosa.*

In order to find the genes encoding the laccases of interest a cDNA library of *S. aeruginosa* was required. For this purpose the λZAP cDNA Synthesis Kit from Stratagene was selected. The kit required significant amounts of clean mRNA, which was obtained by the classical phenol extraction/LiCl precipitation method to isolate total RNA, essentially as described by Chomczynski (Chomczynski and Sacchi, 1987) and subsequently fractionated with oligo-dT coated magnetic beads to obtain messenger RNA. The biomass used for mRNA isolation was selected on

the basis of a time-course experiment measuring change in laccase specific activity over four weeks of cultivation. A flow chart presenting the process of creating a cDNA library using the kit is shown in Figure 5.1. The subsequent Figures 5.2, 5.3 and 5.4 present various steps of this protocol. RNA quality assessment by agarose gel electrophoresis in the presence of formaldehyde (Figure 5.2), strand synthesis controls performed with alkaline agarose gel (kit manual p. 48; Figure 5.3) and the result of cDNA fractionation on Sepharose CL-2B (kit manual p. 22; Figure 5.4) are all illustrated. RNA electrophoresis confirmed that a representative pool of RNA has been isolated, first strand synthesis control showed no problems in the first strand cDNA synthesis (good representation, no hairpins) and the DNA fractionation step allowed to discard fractions smaller that 150 bp, excess of nucleotides and adapters.

The resultant cDNA was used to create two libraries. The first one was a standard λZAP bacteriophage library that could be directly screened either by activity, or nucleic acid hybridisation or converted into an *E. coli* plasmid library by an *in vivo* excision protocol using helper phage, as described in the manual and in the literature (Short et al., 1988). The other library used a customised yeast expression shuttle vector pYES 2.1 to assess the possibility of activity screening of a yeast library for laccase genes.

2.2. Preparation of λZAP library of *S. aeruginosa*.

The cDNA fractions obtained by size exclusion chromatography indicated 1-8 in Figure 5.4 were pooled, ligated to λZAP vector and packaged into bacteriophage lambda with a primary library titer with *E. coli* XL-1 Blue MRF′ of $8x10^5$ pfu ml^{-1}. 97 % of tested clones showed the disruption of the *lacZ′* gene indicating the appropriate cDNA fragment ligation, as shown by plating the phages in the presence of IPTG and X-gal. The primary library was amplified once to obtain a secondary library

titer of $2.5x10^{10}$ pfu ml^{-1}. The secondary library was used in all subsequent screens.

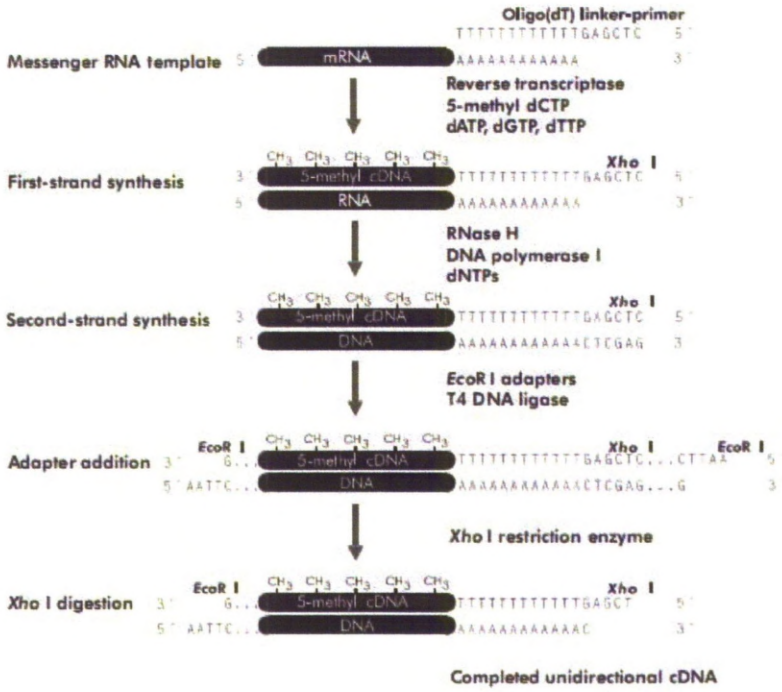


Figure 5.1 The cDNA synthesis protocol, steps prior to ligation to vector. (adapted from λZAP cDNA Synthesis Kit [Stratagene] p.8)

The pool of mRNA is reverse transcribed by reverse transcriptase in the presence of 5-methyl dCTP to create the methylated first strand of cDNA. The reverse transcription is performed with an oligo dT primer that introduces a XhoI site at 3′ end of the resultant DNA-RNA hybrid. The next step involves RNA digestion with RNaseH and synthesizing the second strand of cDNA with DNA polymerase I. When the synthesis is complete EcoRI adapters are ligated to introduce a cloning site at the 5′ end of the double stranded cDNA. Final digestion with XhoI removes the EcoRI site at the 3′ end introduced in the previous step. cDNA methylation prevents internal digestion by XhoI nuclease.

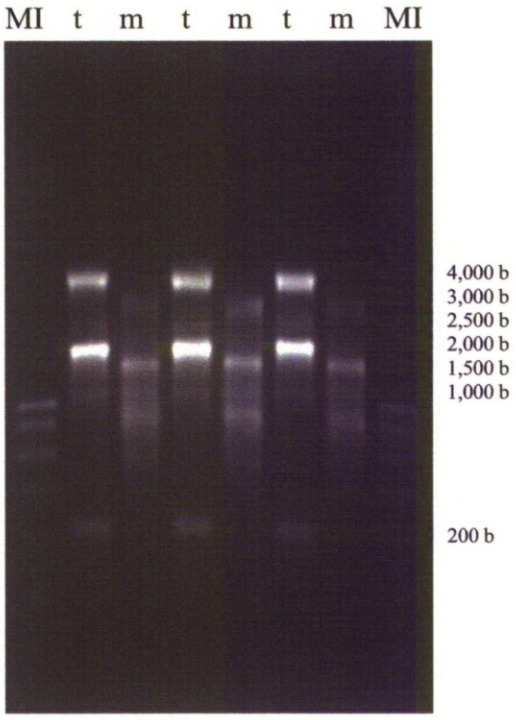


Figure 5.2 *S. aeruginosa* RNA.

The total RNA (t) and mRNA (m) of *S. aeruginosa* separated on a 1.3 % (w/v) agarose gel containing 3.7 % (v/v) formaldehyde; MI – Hyperladder I (range 10000 – 200 bp)

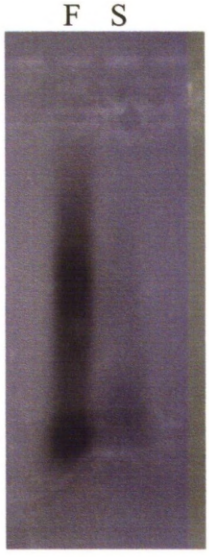


Figure 5.3 *S. aeruginosa* cDNA synthesis controls.
Radiograph showing ^{32}P labelled cDNA synthesis controls of first (F) and second (S) strand run on a 1 % (w/v) alkaline agarose gel. An uniform smear indicates that isolation and synthesis were performed correctly.

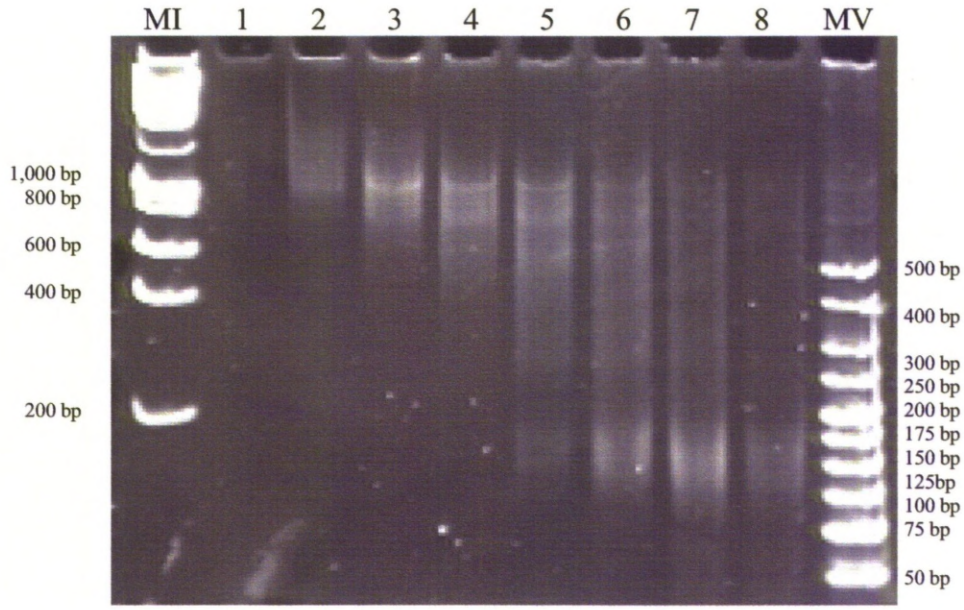


Figure 5.4 *S. aeruginosa* cDNA fractionation.

Fractionation of cDNA on Sepharose CL-2B performed prior to ligation to vector. 16 fractions of 100 µl each were collected and fractions 1-8 presented on this figure were pooled together and used for subsequent analysis. Fractions 9-16 (not shown) were discarded due to the small size of cDNA. 8 µl of those fractions were separated on a 5 % (w/v) polyacrylamide gel to assess the molecular weight of DNA. MI – Hyperladder I (range 10000 – 200 bp); 1-8 – cDNA fractions 1-8 after chromatographic separation; MV – Hyperladder V (range 500 – 25 bp).

2.3. Preparation of plasmid library of *S. aeruginosa*.

The secondary phage library of *S. aeruginosa* was used to create a plasmid library in pBluescript SK-. The *in vitro* excision was performed by co infection of the *E. coli* XL-1 Blue MRF′ strain with the λZAP phage bearing the cDNA inserts and ExAssist helper phage in ratio cells : library phage : helper phage of 10^8 : 10^7 : 10^9. After excision, nonsuppressing *E. coli* SOLR cells were used for transformation and recovery of plasmid DNA. In order to check whether the resultant library was representative the

plasmids were digested with EcoRI/XhoI to release the cDNA from the vector and separated on an agarose gel (Figure 5.5, lane EX Blu).

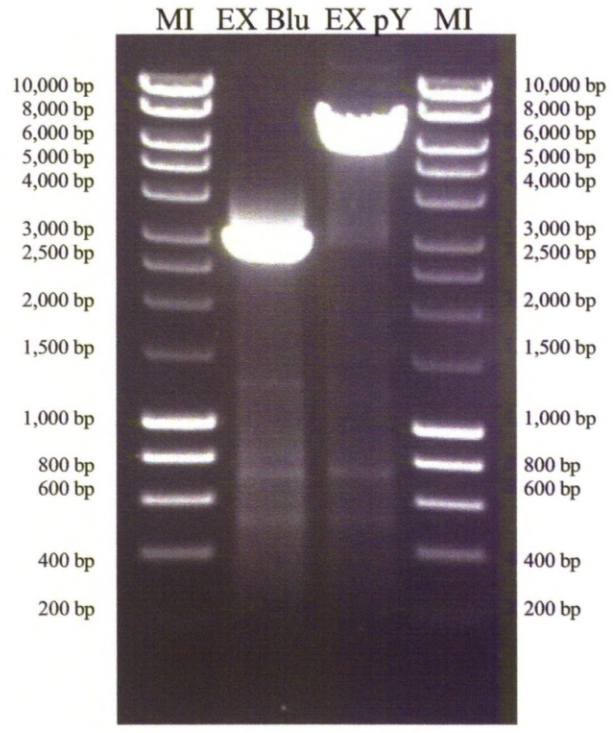


Figure 5.5 Restriction digests of plasmid libraries of *S. aeruginosa.*

Plasmid cDNA libraries were digested with EcoRI and XhoI to release the cDNA from the vectors and run on 1 % (w/v) agarose gel. MI – Hyperladder I (range 10000 – 200 bp); EX Blu – Smear created by EcoRI/XhoI digest of *E. coli* plasmid library in pBlusecript SK- vector; EX pY – EcoRI/XhoI digest of *E. coli – S. ceraviseae* plasmid library in pYES-MCS shuttle vector. The size of an empty vector pBlusecript SK- is 2925 bp; the size of an empty vector pYES-MCS is 5894 bp.

2.4. Preparation of yeast library of *S. aeruginosa*.

2.4.1. Creation of pYES-MCS vector.

To obtain an *S. cerevisiae* vector compatible with the EcoRI/XhoI restriction sites created by the λZAP cDNA Synthesis Kit, it was decided to modify the commercially available pYES 2.1 shuttle vector with a custom made linker made of short complementary oligonucleotide sequences with A overhangs on both sides to utilise the TOPO TA cloning ability of the vector. The two complementary oligonucleotides were annealed to one another and TA-cloned into the vector and the linker orientation was tested by restriction digest and polyacrylamide gel electrophoresis. Figure 5.6 illustrates the process of creating the vector pYES-MCS from those two components.

2.4.2. Introduction of *S. aeruginosa* cDNA to pre-digested pYES-MCS vector.

The unidirectional cDNA created with the kit was ligated into the (pre-digested with EcoRI/XhoI, dephosphorylated) pYES-MCS vector and transformed into the best competent cells available (SoloPack® Gold) to minimise the loss of any information from library. The details of the process are presented in M&M. In order to check if the resulting library was representative the plasmids were digested with EcoRI/XhoI to release the cDNA from the vector and separated on agarose gel (Figure 5.5, lane EX pY). Plasmids were isolated from *E. coli* cells and transformed into yeast for screening purposes with the transformation protocol devised by Gietz (2007).

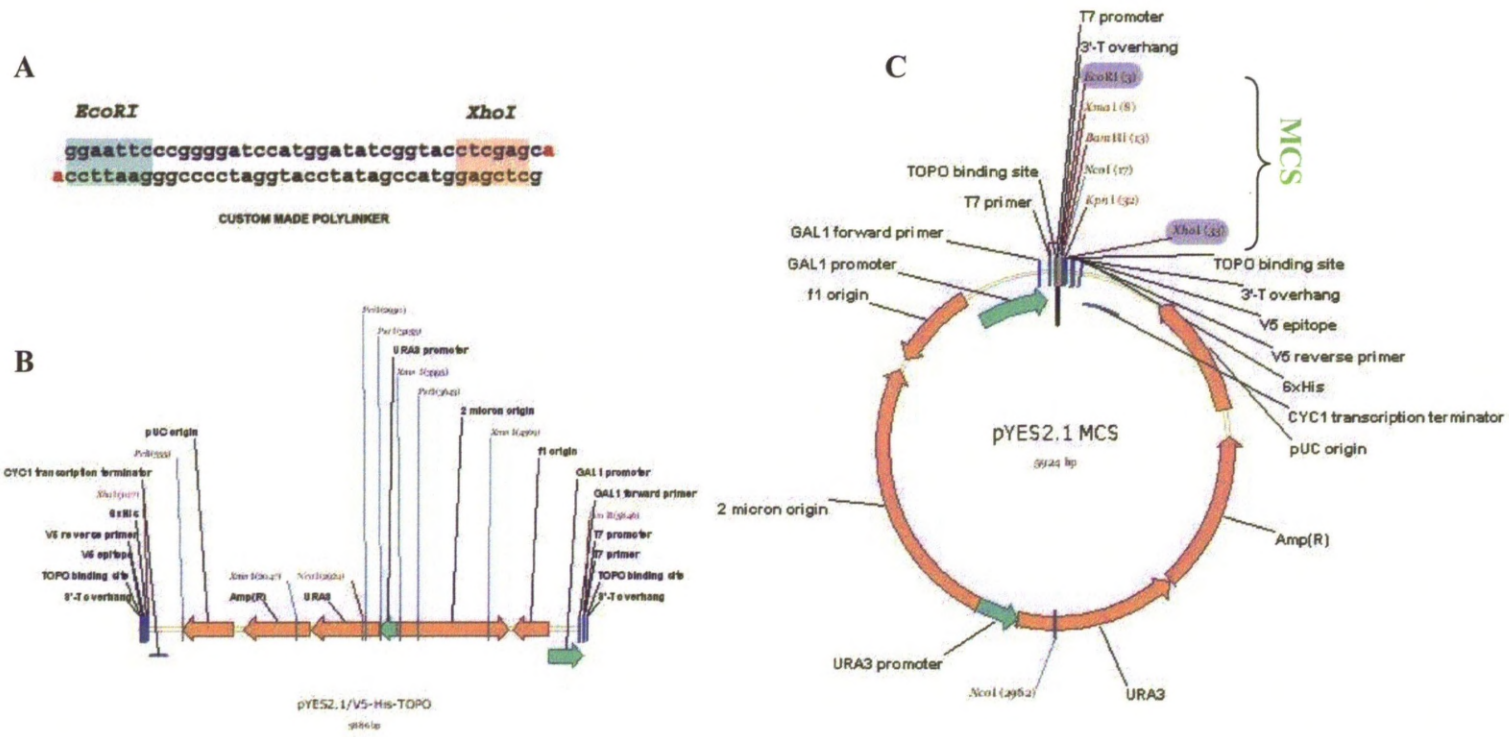


Figure 5.6 Creation of pYES-MCS vector.

A – the sequence of polylinker with TA cloning sites in red and EcoRI and XhoI recognition sites shaded
B –vector pYES 2.1linearised at TA cloning site

C – complete map of circularised vector pYES-MCS with multiple cloning site highlighted in green. EcoRI and XhoI cloning sites shaded in blue.

3. Activity-based screening of cDNA libraries of *S. aeruginosa.*

3.1. Introduction.

In order to screen the cDNA library of *S. aeruginosa* for laccase activity, the *o*-dianisidine assay originally proposed by Antipov (Antipov et al., 2003) and developed by Moore, and the modified ABTS assay proposed by Iffland (Iffland et al., 2001) were used.

3.2. Activity screening of λZAP library o *S. aeruginosa.*

Approximately $5x10^6$ pfu were screened using the adapted *o*-dianisidine assay. In order to allow protein expression the top layer of agar was supplemented with IPTG as an inducer and *o*-dianisidine as an activity indicator. Unfortunately this method did not seem to show any postitve clones despite several modifications of the top layer of agar including pH and the use of ABTS as a chromogenic substrate. Those modifications resulted from advances in characterization of the native proteins of *S. aeruginosa.*

3.3. Activity screening of plasmid library of *S. aeruginosa.*

Unsuccessful attempts to screen the phage library using activity screening prompted questions regarding substrate accessibility and optimal conditions for the reaction. Both colorimetric substrates were used in the reaction buffer in the presence of DMSO as a co-solvent, and further advances in biochemical characterization showed high preference of the enzyme for acidic conditions (pH 3.0). Neither of these features was easy to obtain in an agarose layer due to substrate precipitation and phage lambda viability, respectively. As a result it was decided to use the *E. coli* phagemid library and perform the screen on colony lifts according to the

protocol adapted from Iffland (Iffland et al., 2001). Firstly the phage library was excised *in vivo* from phage lambda by co-infection of host cells *E. coli* XL1-Blue MRF′ with a helper phage. The resultant phagemid library was used to infect an *E. coli* SOLR strain used in activity screening. The process is explained in more detail in Stratagene manual p. 38 and the literature (Short et al., 1988).

The bacteria were plated onto IPTG-containing LA medium to induce protein expression and lifted on nitrocellulose for the colourimetric assay. The nitrocellulose lifts were incubated between ABTS soaked filter papers in optimal conditions of laccase activity found during biochemical characterization of the proteins N and E i.e. McIlvaine buffer pH 3.0 and incubated overnight at 42 °C (Chapter 4). Once again around 10^6 colonies were screened with this method without success; Figure 5.7 illustrates colony lifts and filter paper soaked with chromogenic substrate.

3.4. Activity screening of yeast cDNA library of *S. aeruginosa.*

As a result, two possible conclusions were drawn. Firstly, that *S. aeruginosa* laccase cannot be functionally expressed in *E. coli* and therefore a eukaryotic expression system is needed. Alternatively the presence of a long 5′ UTR in the cDNA sequence cloned into the vector that shifts the start of the codon outside the effective range of the *lac* promoter. As a result it was decided to try to repeat the experiments with *S. cerevisiae* as a new host for the library. Unlike *E.coli* cells, yeast cannot be directly plated onto expression medium due to a very low survival rate on galactose. To address this issue the whole library was plated onto minimal growth medium and subsequently replica plated onto a galactose-containing expression medium. The trials using both methods previously tested with *E. coli* libraries i.e. agarose overlays with ABTS and colony lifts incubated in the presence of ABTS-soaked filter papers were performed and proved equally unsuccessful.

As a result of these negative screening trials, it was decided to shift attention to molecular screening and the use of degenerate primers/probes to screen the libraries of *S. aeruginosa.*

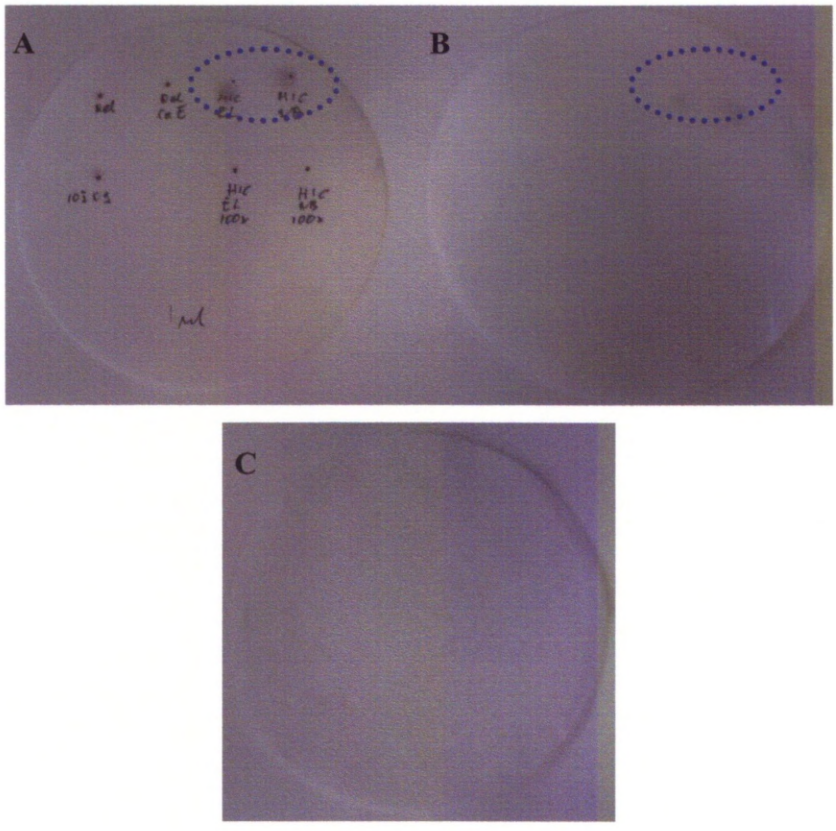


Figure 5.7 Colourimetric lift screen assay of a plasmid cDNA library of *S. aeruginosa.*

Filter papers were soaked in McIlvaine buffer pH 3.0 containing 10 mM ABTS and sandwiched with colony lifts wrapped in cling film. Green product of ABTS oxidation is indicated by green spots and highlighted in blue. A – Control filter with Phenyl Toyopearl fraction of native laccases N and E spotted on the membrane (1 µl of pooled protein prior to concentration); B – The chromogenic response of ABTS on the filter paper of purified laccases; C – The lack of chromogenic response of ABTS on the filter paper during assay conditions

4. Molecular-based screening of cDNA libraries of *S. aeruginosa.*

4.1. Degenerate primer PCR screening of *S. aeruginosa* cDNA libraries.

The protein sequences obtained from mass spectrometric sequencing of spots from 2D gels (Chapter 4) were used to design degenerate primers to amplify parts of genes from the cDNA library, which could subsequently be used as long probes to probe the phage library to identify full laccase sequences. Unfortunately almost all the protein sequences (Chapter 4) contained a high content of leucine/isoleucine and lysine/glutamine. These amino acid pairs have the same molecular weight making it impossible to distinguish which amino acid was present in a peptide from sequencing.

Consequently, given the degeneracy of the genetic code the leucine/isoleucine pair could be encoded by up to nine separate codons, whereas lysine/glutamine pair can be encoded by a reasonable four different codons and a general formula of MAR (aMino Adenine puRine) in IUPAC notation. On the other hand tryptophan and methionine are coded by a single nucleotide triplet, but these amino acids were not abundant among the selected peptides (just one tryptophan). When dealing with a well known model organism a codon bias table can be used to decrease the degeneracy of a primer by implementing only the most common codons for a particular organism (Balland et al., 1985). However since *S. aeruginosa* is an organism with very little genetic information available, the use of a codon bias table was not possible. To address the problem of too high level of probe degeneracy, it was decided to include deoxyinosine (I) for most of undetermined nucleotides (N) obtained from reverse translation of protein sequences. Deoxyinosine is a non-standard nucleotide that is capable of pairing with all standard nucleotides in a double-stranded DNA (Ohtsuka et al., 1985). Therefore each substitution

of N with I decreases the degeneracy of the primer fourfold. Deoxyinosine-containing primers were found to work with *Taq* polymerase but not proofreading *Pfu* due to the 3′→5′ exonuclease activity of the latter (Knittel and Picard, 1993).

Out of all the protein sequences obtained from the proteomic analysis described in details in Chapter 4, three were selected for PCR-based screening. Table 5.1 summarises all the properties of chosen sequences and corresponding primers.

Table 5.1. Peptide sequences and their reverse translation in primer design. DG – degenaracy of the primer

Peptide sequence	Selected sequence	Primer name and orientation	Primer sequence	DG
GPNANLFLQNK	GPNAN	GPNAN_REV	TAA ICG TAA NCC NG	16
VVNATAV	VVNATAV	VVNATAV_FWD	GTN GTI AAT GCN ACI GC	16
KQVWNVIV	VWNVIV	VWNVIV_REV	TG CTA ITG TAA GGT ITG	1

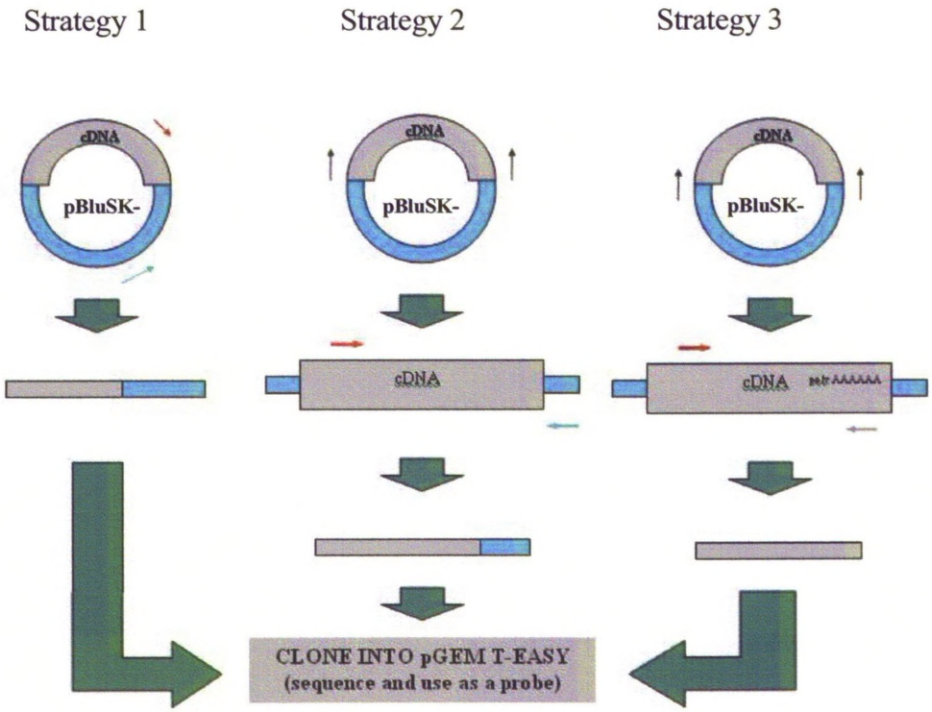
To amplify the probe, PCR using the cDNA plasmid library created in this Chapter (point 2.3) as a template was performed. Several combinations of primer pairs were tested and are summarized in Figure 5.8.

- The first strategy involved using one gene specific degenerate primer (Table 5.1) in conjunction with a pBluescriptSK(-)-specific primer, the name of each of these primers correspond to the name of the degenerate primer i.e. primer based on protein sequence GPNAN had a degenerate primer GPNAN_REV (Table 5.1) and a primer that was specific to a site in pBluescriptSK(-) vector and had similar binding characteristics to GPNAN_REV. It was called GPNAN_FWD. Analogous primer pairs were designed for peptide sequences VVNATAV and VWNVIV (Table 5.1, Figure5.8)
- The second strategy required two amplifications. Firstly the cDNA inserts were amplified with primer pair InsAmp_FWD and

InsAmp_REV to create another template. These primers were designed to amplify whole pool of cDNA from the library as binding sites were selected approximately 250 bp from the cDNA insert. This product was then purified and used in second PCR with specific primers. One of the primers was a degenerate primer from Table 5.1, another one was either InsAmp_FWD or InsAmp_REV.

- The third strategy required two amplifications. Firstly the cDNA inserts were amplified with primer pair InsAmp_FWD and InsAmp_REV to create another template. These primers were designed to amplify whole pool of cDNA from the library as binding sites were selected approximately 250 bp from the cDNA insert. This product was then purified and used in second PCR. For the reaction VVNATAV_F primer and poly dT primer were used.

Unfortunately none of the tested combinations of primers gave any interesting products. The combinations of degenerate primers were also tested as well as different modifications of PCR mix and cycling conditions. As a result of these amplifications either no product at all was obtained or some artefactual primer independent products were visible on an agarose gel (data not shown). A few amplification products were cloned, sequenced and proved to be misamplifications.



PCR primer pairs					
Strategy 1		Strategy 2		Strategy 3	
1st PCR	GPNAN_FWD, GPNAN_REV, VVNATAV_FWD, VVNATAV_REV, VWNVIV_FWD VWNVIV_REV	1st PCR	InsAmp_FWD InsAmp_REV	1st PCR	InsAmp_FWD InsAmp_REV
		2nd PCR	InsAmp_FWD VWNVIV_REV InsAmp_FWD GPNAN_REV VVNATAV_FWD InsAmp_REV	2nd PCR	VVNATAV_FWD polyT_REV

Figure 5.8 Summary of strategies and for PCR probe amplification.

Scheme at the top presents an overview of primer binding sites for each strategy, the table below indicates primer pairs used for each amplification. Each different primer pair is highlighted in the same colour.

4.2. Nucleic acid hybridisation-based screening of *S. aeruginosa* cDNA libraries.

Lack of success in amplifying the gene fragments of *S. aeruginosa* laccases prompted us to use the last available option of screening the cDNA library *via* plaque hybridization using degenerate primers. In order to do that, the degenerate oligonucleotides were 5′ labelled with ^{32}P and initially used in the screening process against a Southern blot made of cDNA insert amplified with InsAmp_FWD and InsAmp_REV primers as described before. This pre screening was used to optimize the hybridization conditions for actual plaque screening. The cDNA was cleaned up with PCR clean up kit and half was digested overnight with EcoRI and XhoI to liberate the real cDNA fragment from two 250 bp long pBluescript fragments prior to electrophoretic separation. The agarose gel and the blot is presented in the Figure 5.9. Blotting efficiency was demonstrated by hybridizing ^{32}P labelled universal oligo dT primer to cDNA of *S. aeruginosa*.

Optimization of hybridisation conditions was performed in the following manner. The initial hybridisation was performed at 37 °C to allow even nonspecific probe binding (due to the degeneracy of probes and their very short length). After hybridisation the membrane was washed at 37 °C first and autoradiogram showing hybridization pattern was scanned. During the process of hybridization, it was found that only the probe based on the peptide sequence VVNATAV showed any specificity towards cDNA of *S. aeruginosa*. To obtain more reliable results the temperatures of washing were gradually increased. At the temperature of 48.5 °C the last clear image was observed. Above this temperature all signal was lost due to probe dissociation from the blot. The final result of this optimisation is presented in Figure 5.10. The satisfactory hybridization signal was highlighted in red and overlaid with the gel picture.

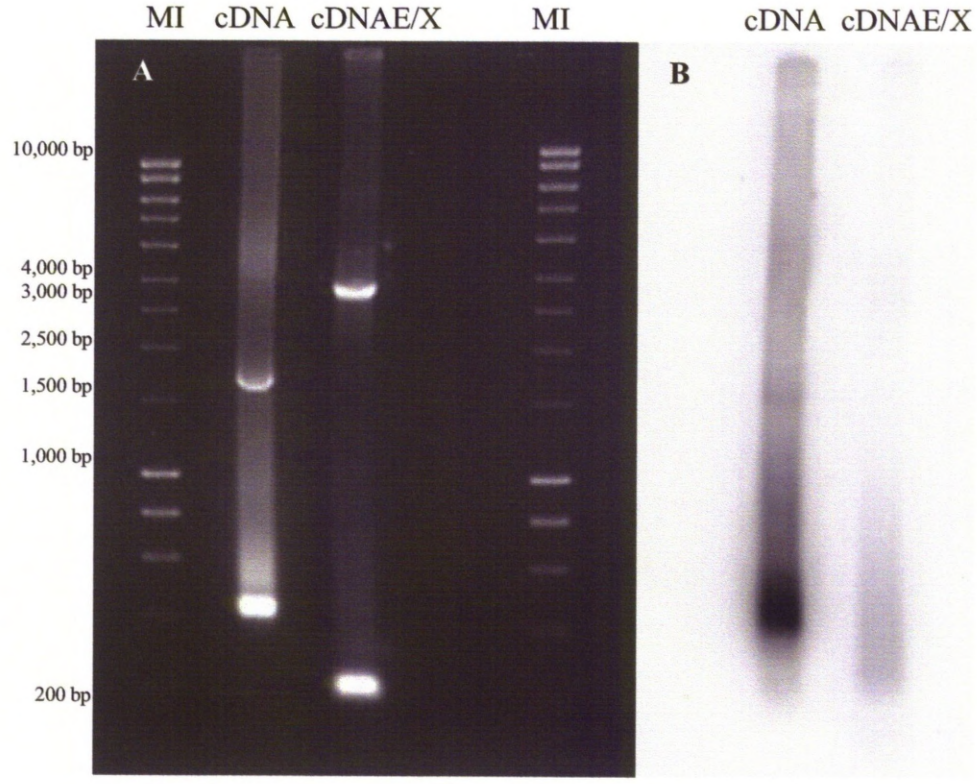


Figure 5.9 The separation and blotting of cDNA insert of *S. aeruginosa*.

The cDNA of *S. aeruginosa* was separated on 1 % (w/v) agarose gel and blotted.

A – electrophoretic separation of PCR amplified insert of *S. aeruginosa* plasmid cDNA library: MI – Hyperladder I (range 10000 – 200 bp); cDNA – non digested cDNA separated on gel; cDNA E/X– EcoRI/XhoI digested cDNA separated on gel

B – autoradiogram from probing the membrane with ^{32}P labelled universal oligo dT primer; cDNA – non digested cDNA separated on gel and blotted; cDNA E/X– EcoRI/XhoI digested cDNA separated on gel and blotted

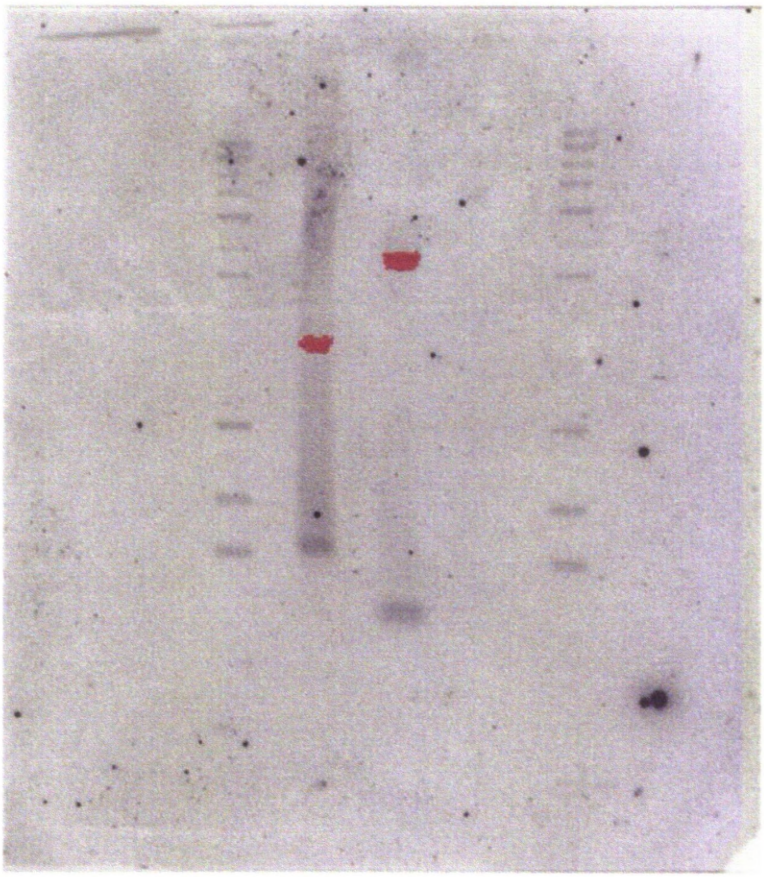


Figure 5.10 Optimisation of hybridization conditions for VVNATAV probe.

The hybridisation signal at the temperature T = 48.5 °C was marked in red and overlaid onto the agarose gel picture from Figure 5.9.

On the basis of these results the experiments of plaque hybridization using the same probe and hybridization conditions were performed. Around 10^6 pfu were initially screened yielding over 200 positive hybridization signals (Figure 5.11), which were then plated at lower density (less than 500 pfu per plate to assure single plaque isolation) and subjected to hybridization once again.

The hybridisation experiment showed 107 positive plaques, which where extracted from the plate. The phagemids were excised (library manual p. 39) and subsequently isolated from *E. coli* SOLR cells. All 107 phagemids were digested with EcoRI and XhoI to check the size of the

inserts on an agarose gel, which was then blotted. The blots were subjected to hybridisation with ^{32}P labelled universal oligo dT probe to check that inserts contained full length non degraded cDNA and highlight which band may contain the insert of interest. The cDNA bands of a length over 1 kb and showing positive hybridisation with universal oligo dT probe were considered for sequencing. The gel and result of hybridisation with dT probe are presented in Figures 5.12 and 5.13 respectively.

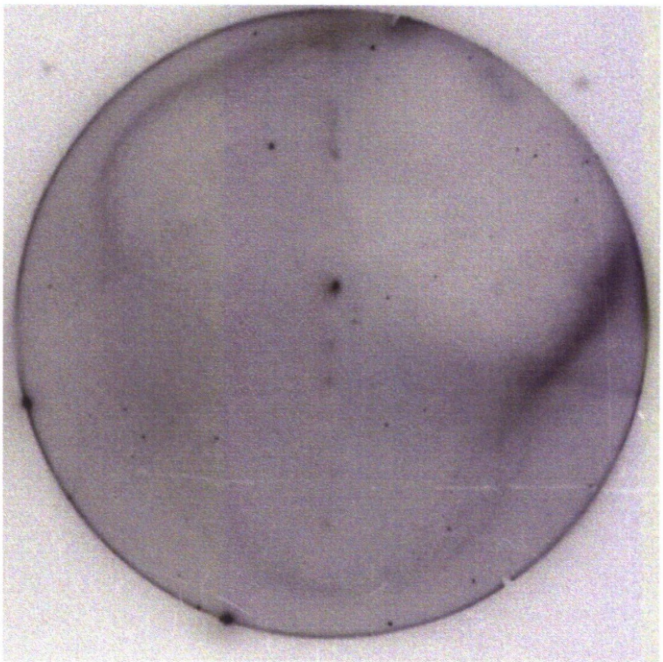


Figure 5.11 Typical hybridization signals with VVNATAV probe of during the phage library screening.

A typical hybridisation signal of a membrane used in plaque screening. Membrane was washed 48.5 °C before scanning. Plaques indicated by this screen were isolated; each clone was converted into plasmid and used in subsequent rescreening.

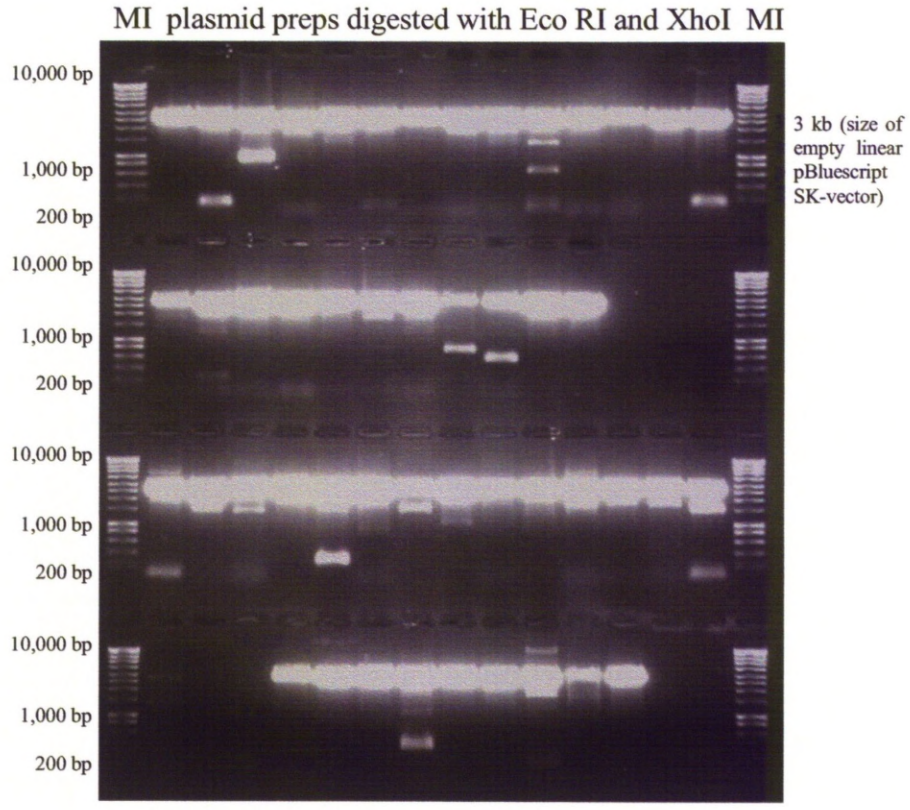


Figure 5.12 Positive plaques rescreening.

Plaques indicated by this screen were isolated; each clone was converted into plasmid. Plasmids were digested with EcoRI/XhoI to release the insert. MI – Hyperladder I (range 10000 – 200 bp); plasmid preps digested with Eco RI and XhoI – plasmid preps prepared from overnight cultures of SOLR cells infected with phagemids indicated by probe hybridization experiments; 3 kbp. The linear size of pBluescriptSK(-) vector digested with EcoRI and XhoI

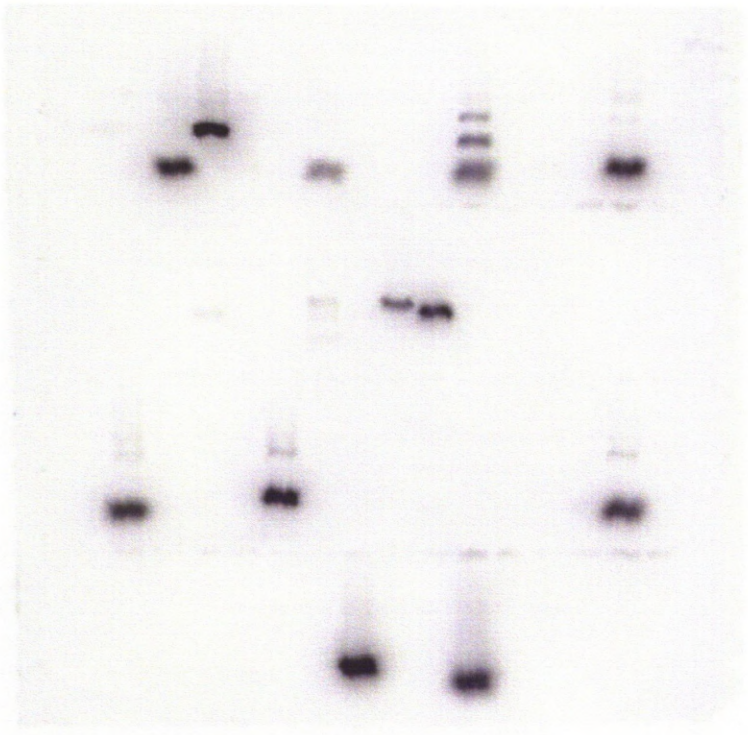


Figure 5.13 Autoradiograph of Figure 5.12.

Plasmids isolated from plaque hybridization screen digested with EcoRI/XhoI after blotting and hybridisation with oligo dT probe.

Only six plasmids fulfilled those two criteria and were sent for sequencing, unfortunately none of these sequences showed any similarity to laccases when analysed with BLAST. None of these sequences matched any proteomic data either. As a result those sequences must be regarded as false positives caused by lack of specificity of the probe due to its very short length.

5. Summary.

In a pursuit of the genes encoding two laccases characterised in Chapter 4 three cDNA libraries of *S. aeruginosa* were constructed. The libraries in bacteriophage lambda, bacterial plasmid vector and bacteria-yeast shuttle vector were screened using an array of methods. Those methods can be divided into two main categories – based on the detection of protein activity *via* its reaction with a chromogenic substrate and molecular techniques based on base pairing between oligonucleotide primers/probes designed from known protein sequence obtained through mass spectrometry. None of the methods proved satisfactory. No results were obtained through protein activity-based screens, probably due to no or very low expression of the proteins. This might have been caused by inability to express the functional protein by either *E. coli* or *S. cerevisiae* or due to incompatibility of the genetic insert with the functional range of either lac or GAL promoter. The complete lack of results with the functional screen prompted us to try molecular methods, the first method was intended to use PCR and degenerate primers to find a fragment of a gene of a laccase of interest and use it to screen the lambda ZAP cDNA library of *S. aeruginosa* to find a full length sequence. Unfortunately this theoretically promising route failed probably due to lack of specificity of the degenerate primers and we turned to the last resort i.e. to try to hybridise short degenerate oligonucleotide to a complex library. Significant signal background was expected; therefore the process of optimising the hybridisation conditions was performed with a Southern blot of cDNA to find the optimal probe hybridization conditions. Unfortunately this method did not prove successful and yielded many false positives. This failure prompted us to look for non conventional methods to find the desired sequences and turned our attention to next generation sequencing methods, namely 454 pyrosequencing developed by Roche and routinely used in the Centre for Genomic Research of the University of Liverpool.

Chapter 6

THE USE OF 454 SEQUENCING IN A PURSUIT OF *S. AERUGINOSA* LACCASE GENES

1. Introduction.

The failure of traditional attempts to find the laccase genes and arrival of Mr. Tomasz Florczak in the lab with the task of finding the gene of a lipase from a non-model organism *Geomyces sp. P7,* prompted us to look for alternatives to identify these sequences. Our attention turned to the possibility of using next generation high throughput sequencing. It was suggested that both the sequences of the laccase cDNA and the lipase cDNA could rapidly be identified by high throughput sequencing of the transcriptome of the two organisms generated when enzyme activities were at their peaks. At the time, two machines were available in the Centre for Genomic Research of Liverpool University, a 454 GS FLX Titanium pyrosequencer developed by Roche and a SOLiD ligation sequencer produced by Applied Biosystems. The recently developed Titanium pyrosequencing system seemed to be an interesting option as it generates much longer reads than the SOLiD system (400 bp vs. 35 bp) at lower costs per run. The longer reads were absolutely essential when choosing the sequencing method as financial constraints allowed only very low coverage sequencing. It is also commonly known that SOLiD is superior in resequencing of known genomes rather than assembling new ones, where longer reads are necessary (Metzker, 2010). However, the longer read lengths come at a cost; the 454 system is almost perfect when sequencing different bases one after another, but homopolymers present a significant problem, due to the improper assigning of number of bases at a given light

signal which then causes frame shifts in the sequence. To overcome this problem resequencing of the sample is highly recommended, preferably with a more accurate system like SOLiD or Genome Analyser from Illumina. Unfortunately this method was not possible due to financial constraints.

This chapter focuses on work done on the *S. aeruginosa* transcriptome only, an identical approach was applied to *Geomyces sp. P7* lipase. Results of this collaborative work are currently in preparation for publication (Florczak, Daroch et al.).

2. Creation of cDNA of *S. aeruginosa* for high throughput sequencing.

The sequencing process requires around 1 μg of high quality normalized cDNA. The mRNA was isolated from *S. aeruginosa* after 4 weeks of growth using the methodology described previously. The cDNA was prepared with a SMART™ PCR cDNA Synthesis Kit recommended by James Hartwell. Hartwell's group has successfully used this particular kit for their 454 transcriptome work. Firstly it was attempted to optimize the PCR cycling conditions to avoid excessive overrepresentation of the most abundant sequences due to the exponential character of the PCR reaction. The result of cycle numbers optimisation required by the manufacturer's protocol (p. 19) is presented in Figure 6.1. The appropriate number of PCR cycles giving a representative cDNA pool was estimated to be 20 cycles. This evaluation was performed on the basis of gel electrophoresis and the manufacturer's instruction to use "one cycle fewer than needed to reach a plateau".

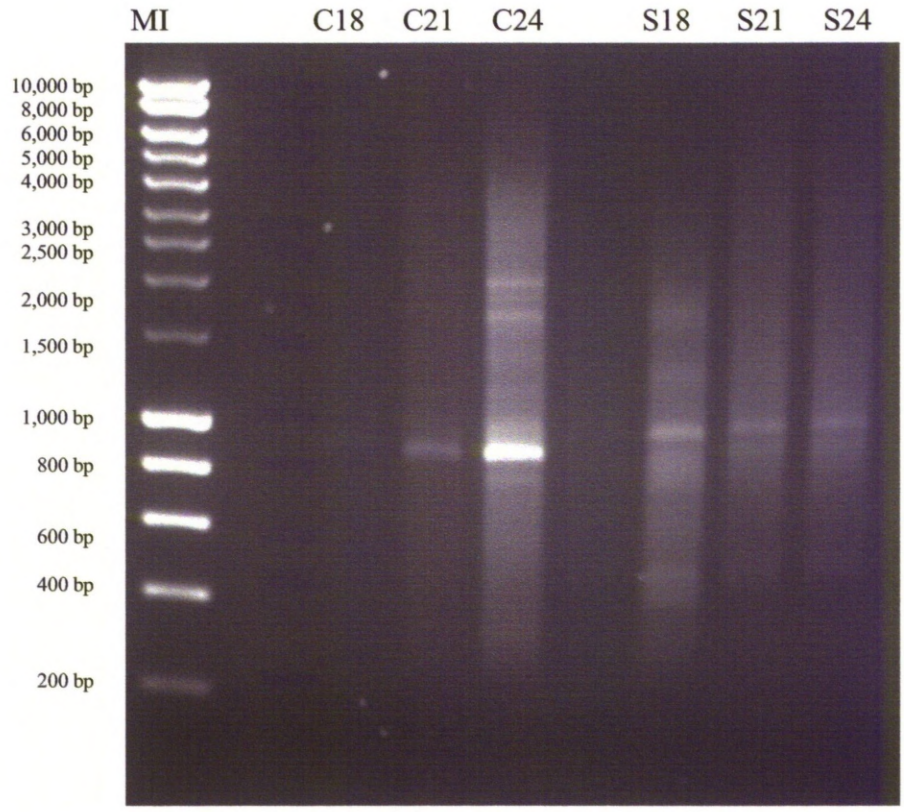


Figure 6.1 Optimisation of PCR cycling conditions for cDNA synthesis.

Optimisation of PCR cycling conditions to create representative cDNA pool of *S. aeruginosa* for high throughput sequencing was performed according to manufacturer's guidelines. MI – Hyperladder I (range 10000 – 200 bp); C18 – 18 cycles with control placental cDNA; C21 – 21 cycles with control placental cDNA; C24 – 24 cycles with control placental cDNA; S18 – 18 cycles with *S. aeruginosa* cDNA; S21 – 21 cycles with *S. aeruginosa* cDNA; S24 – 24 cycles with *S. aeruginosa* cDNA.

The SMART cDNAs were given to Dr Margaret Hughes from the Centre for Genomic Research of the University of Liverpool who then performed the transcriptome sequencing using one sixteenth of a 454 sequencing plate for the library. This scale of sequencing should result in approximately twenty five million base reads per library (Hughes, personal communication), which we assumed should be enough to provide us with the data we needed.

3. The output from the 454 sequencer.

When the sequencing process was completed the basic bioinformatic work associated with presenting it in an accessible manner was performed by Dr Kevin Ashelford form the Centre for Genomic Research of the University of Liverpool. The raw data were screened with a filter that rejected the poor quality sequences and the rest of the data were then assembled to form contigs. Table 6.1 summarises the data obtained from the sequencing process both in terms of the amount and characteristics of the raw data and the assembled contigs. The amount of data output from the sequencer was a little bit short of what had been expected, however it gave enough data to have a general view on the transcriptome of *S. aeruginosa*. After assembly 4,420 contigs were generated from 62,907 singletons (single non assembled read outs). All parameters of the sequencing read are summarised in Table 6.1

Table 6.1 The output from 454 sequencer.

Feature	***S. aeruginosa***
raw data file	MID2_FPDG3U302.fas
no. of bases	19,991,305
no. of sequences	62,907
mean read length	317.79
median read length	379.00
contig data file	454AllContigsSA.fna
no. of bases	2,045,330
no. of sequences	4,420
mean read length	462.74
median read length	408.00

4. The automated annotation of transcriptome with MG-RAST.

Our first attempt to annotate the transcriptomes of both organisms was to use an automated annotation server MG-RAST (Meyer et al., 2008) which is a tool used for metagenome annotation, i.e. traditionally is used for annotating data from non-model sources, unlike some other annotation tools like TACT (Yamasaki et al., 2006). This online tool deals only with model organisms and compares new data with databases of such organisms as *Homo sapiens*, *Arabidopsis thaliana*, *Drosophila melanogaster*, and does not allow customizing the search for a fungal database.

The MG-RAST server presented the statistical output regarding the basic properties of the contigs sequences, such as their length (Figure 6.2), distribution and GC content.

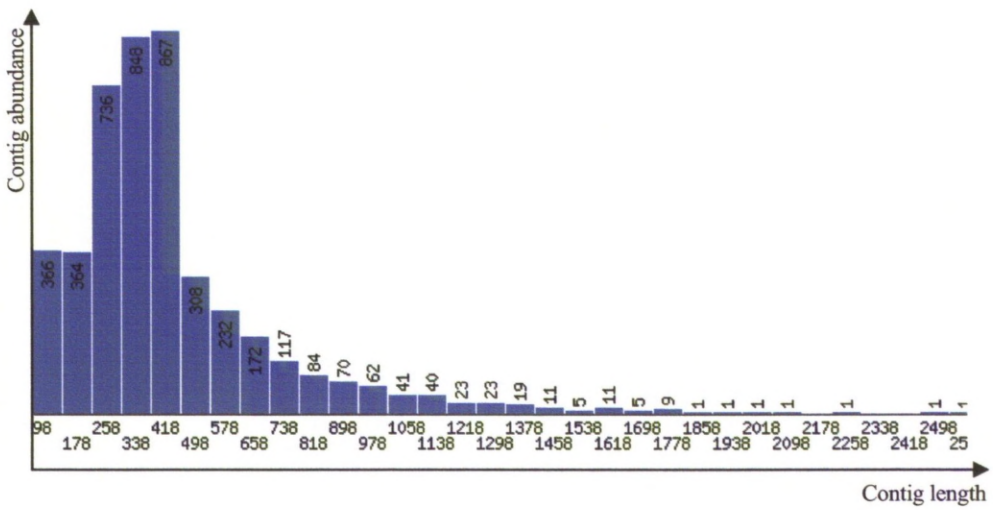


Figure 6.2 The assembled contig length distribution of *S. aeruginosa* cDNA sequenced with 454 pyrosequencer.

Contigs assembled from the sequencing read are relatively short with the vast majority of the contigs ranging 250-400 bp. Very few contigs can contain the full length sequence of a laccase.

The small scale of sequencing resulted in an average contig length of 463 bp which is short of the expected transcript length of above 1500 bp characteristic of fungal laccases (Rodriguez et al., 2008, Bleve et al., 2008). The contig distribution curve in Figure 6.2 indicated that there was only a small possibility that the full sequence of a laccase could be found as only 37 contigs were longer than 1500 bp. The vast majority of the contigs were no longer than four to five hundred base pairs. Therefore it was assumed that once a contig with a fragment of laccase is identified, another method would have to be used to obtain full length sequence. The automatic annotation process did not give us the results we expected i.e. complete annotation of contig sequences. Despite applying the least stringent annotation criteria available (the E value of 0.01), the MG-RAST algorithm was only able to annotate 26.7 % of *S. aeruginosa* sequences. This analysis of the basidiomycete's transcriptome indicated no laccases annotated by the automated server. As a result it was attempted to annotate contigs using a local release of BLAST.

5. Local BLASTX as a tool to search for genes of interest from contigs generated by 454 pyrosequencing.

The very low annotation rate (26.7 %) of *S. aeruginosa* contigs by MG-RAST suggested that an alternative method that will annotate all the sequences must be implemented. To do so, a local BLAST was used. Establishment of a local BLAST program on a standalone PC is quite easy and allows much higher throughput than a web interface provided by NCBI or EMBL. To run the analysis, several preparatory steps were required. First the DNA and protein databases were downloaded from NCBI FTP server, secondly a test run with 100 contig sequences was performed to find the most appropriate algorithm for the annotation. Two different algorithms were tested; both require nucleotide sequence as an input; however they use different reference databases for the analysis.

BLASTN compares nucleotide sequence with nucleotide database, whereas BLASTX translates the DNA sequence into protein sequence in three different reading frames and compares them with the protein database. Those preliminary data suggested that BLASTX gives much more data of better quality. This was due to two main reasons; firstly the codon preferences of different organisms differ significantly, therefore protein sequence similarity does not have to reflect nucleotide similarity. Secondly the homopolymer problem of 454 sequencing causes frame shifts that disrupt nucleotide sequence comparisons (there is a big penalty for indels in nucleotide search), however they have no effect for a translated nucleotide sequence that is compared in three different reading frames anyway. The test results were compared with the web interface of BLASTX and found satisfactory.

Once the decision to use BLASTX was made, the analysis of all 4,420 contigs was performed on a standard desktop PC; it took several days and resulted in 522 MB text file of sequence similarity comparisons.

The text file was then searched for a few key terms to highlight the contigs that are to be used for subsequent analysis. Those terms were: "laccase" and "oxidase". The term "laccase" will highlight any similarity of a contig to a laccase, whereas "oxidase" highlights similarities to all oxidases including multicopper oxidases, peroxidases and haloperoxidases. After this first selection the contigs were tested again with BLASTX, this time using the web interface that also allows domain similarity searches to be performed. The data were narrowed to nine contigs, all of them summarized in Table 6.2. It is worth mentioning that *S. aeruginosa* contig 1192 showed some sequence similiarity to the putative haloperoxidase of *Caldariomycetes fumago* (E value of 4e-05) and very high similiarity to the *A. bisporus* putative chloroperoxidase that Moore unsuccessfully attempted to clone (Moore, 2007). These results gave us the indication that an elusive haloperoxidase reported by Moore could be identified and cloned.

Table 6.2 Summary of all contigs selected from local BLAST X analysis.

Contig	Length	E value (best meaningful)	Best meaningful homolog*
502	433	2e-20	Lac2 protein [Phlebia radiata]
518	342	2e-09	laccase [Cyathus bulleri]
641	604	1e-57	laccase 3 [Coprinopsis cinerea]
968	537	9e-49	laccase [Cyathus bulleri]
1153	457	1e-67	laccase [Pholiota nameko]
1577	321	2e-33	laccase [Schizophyllum commune]
3441	800	2e-65	laccase [Laccaria bicolor S238N-H82]
4378	672	2e-83	laccase [Pholiota nameko]
1192	604	1e-59	putative chloroperoxidase [Agaricus bisporus]

*By meaningful it is meant a homologue that indicates the type of a protein. Annotations: "predicted protein", "unnamed protein product" and "hypothetical proteins" without classification are not regarded as meaningful.

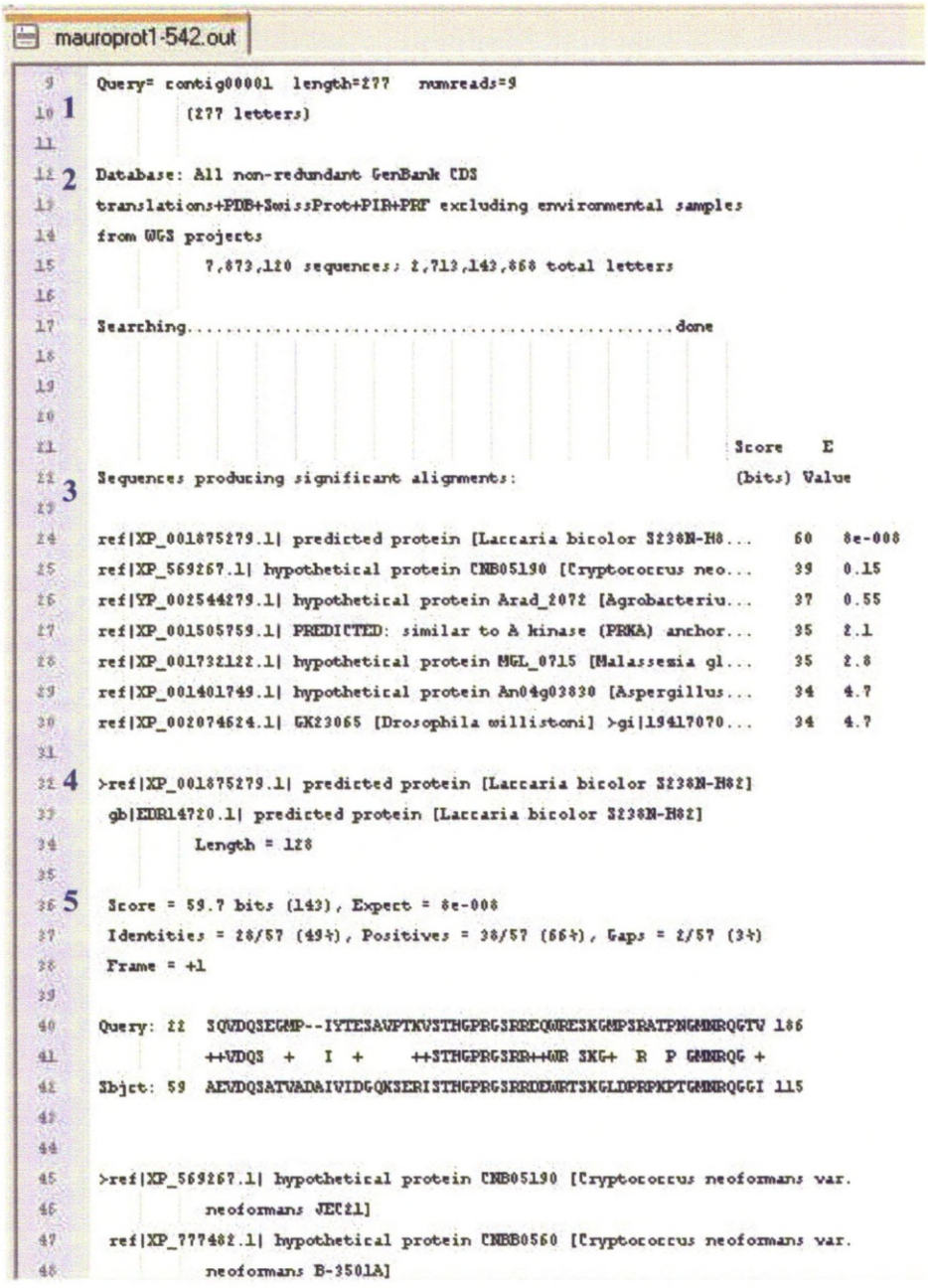


Figure 6.4 A typical output of local BLASTX search.

Similarity search analysis was performed with local BLASTX on *S. aeruginosa* contigs created with 454 sequencing. 1. Query – name of the contig, its length, and number of sequence reads; 2. Database – the type of database; 3. – Highest score list of sequence comparison; 4. – Detailed overview of sequence comparison; 5. – Detailed overview of score and sequence alignment

The analysis resulted in highlighting contigs containing eight potential genes of laccases and a potential haloperoxidase. The length of these contigs ranged from 243 to 800 bp, the values much short of desired 1500 bp of a full length sequence we expected, but long and accurate enough to design nondegenerate primers to amplify our genes of interest from plasmid cDNA library of *S. aeruginosa*.

6. Full-length gene amplification by inverse touchdown PCR and gene cloning.

In order to obtain the full length DNA sequences of the laccase genes it was decided to use the sequence information provided by selected contigs to amplify the missing parts of the gene *via* inverse touchdown PCR on plasmid cDNA libraries of *S. aeruginosa* created earlier. This technique is a combination of inverse PCR (Ochman et al., 1988) that is used to amplify circular templates outside the region of known sequence and touchdown PCR (Don et al., 1991) that implies the stepwise decrease of annealing temperature to avoid the need of careful optimisation of the cycling conditions and to increase the specificity of the product at the same time. Due to the complexity of the procedure we decided to further decrease the unspecific primer binding by application of PCR enhancer that destabilizes mismatched primer-template complexes (Perfect Match PCR Enhancer manual p.2). The nine selected contigs were once again analysed with regard to their accuracy and compared with the singletons from which they were assembled to analyse which parts of the sequence were most accurate. Inverse primers were designed to match only the most certain regions of the contigs and it was attempted to avoid homopolymer regions that are problematic when using 454 technology. Plasmid cDNA libraries were used as templates for this reaction to facilitate the recovery of the product and to allow straightforward sequencing. The amplification was designed to amplify whole plasmid sequence. Out of all amplification

products the focus was only on the PCR products larger than 3 kb (the size of pBluescript SK(-) plasmid) plus the size of the insert i.e. at least 1 kb. The concept is summarized in Figure 6.5. Once the product was obtained it was cut out from the gel, purified and recircularized to obtain a functional plasmid. It was subsequently transformed into *E. coli* cells to multiply. Successful transformants were then grown overnight to isolate plasmids in quantities sufficient for sequencing with standard M13 sequencing primers compatibile with the pBluescriptSK(-) plasmid to identify the missing parts of laccase sequence. Table 6.3 presents the names of primers used for the amplification, Figure 6.6 their sites and Figure 6.7 shows a typical gel prior to band extraction. In cases when more than one product of appropriate size was obtained the bands were cut out separately and subsequently processed. The horizontal dashed line represents the size of pBluescript SK(-) vector i.e. 3 kb. The desired PCR product must be larger than 3 kb all lower bands were discarded as unspecific products of amplification.

Table 6.3 Summary of PCR primers used in inverse touchdown PCR.

Contig	**FWD primer**	**REV primer**
502	Mauro502IvP_F	Mauro502IvP_R
518	Mauro518IvP_F	Mauro518IvP_R
641	Mauro641IvP_F	Mauro641IvP_R
968	Mauro968IvP_F	Mauro968IvP_R
1153	Mauro1153IvP_F	Mauro1153IvP_R
1577	Mauro1577IvP_F	Mauro1577IvP_R
3441	Mauro3441IvP_F	Mauro3441IvP_R
4378	Mauro4378IvP_F	Mauro4378IvP_R
1192	MauroH1192IvP_F	MauroH1192IvP_R

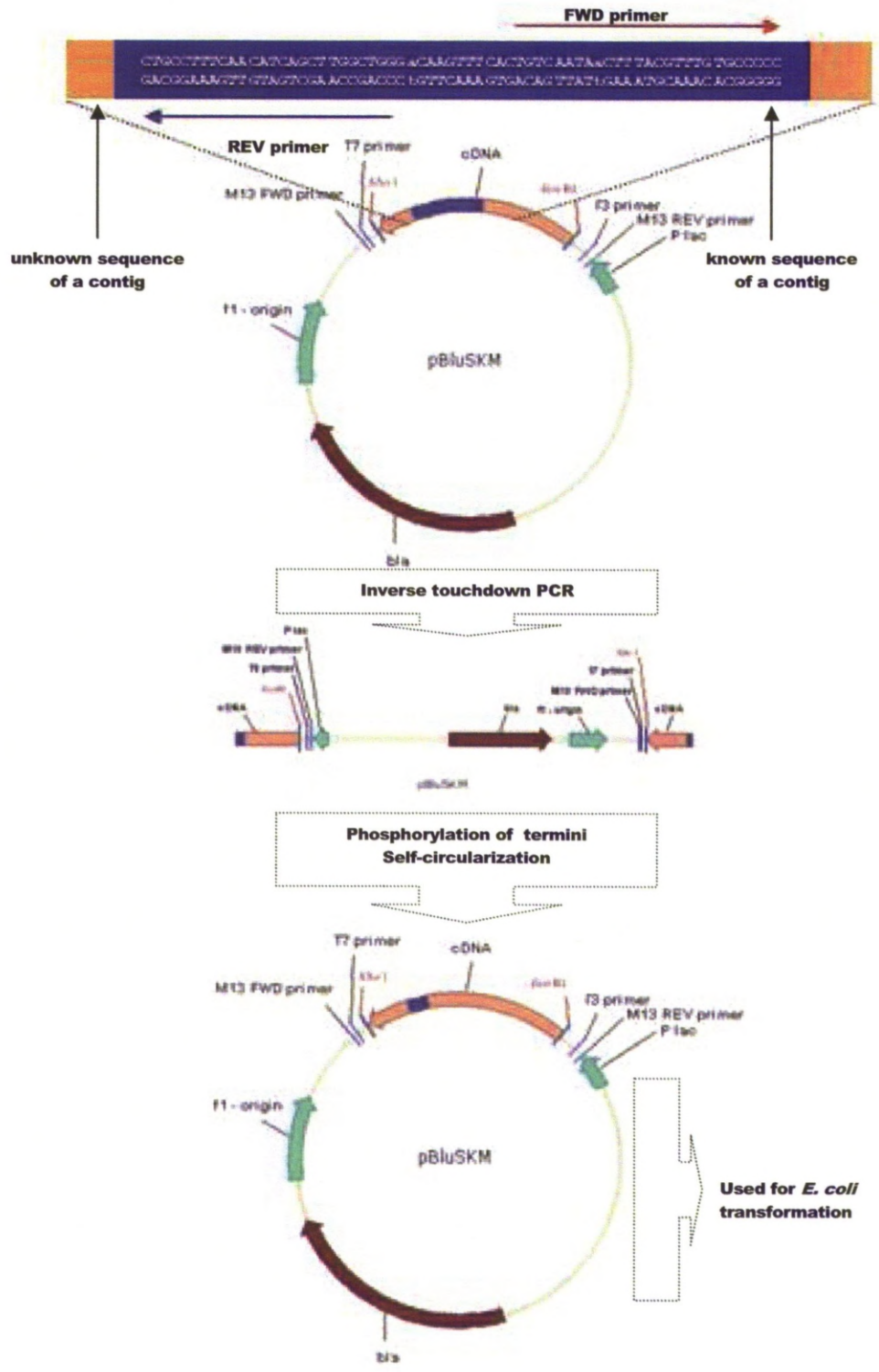


Figure 6.5 The method of amplifying the missing parts of the genes *via* inverse touchdown PCR.

>contig00502 length=433 numreads=4

taTTTCATTcGCGCTGTTCCTAAcATTGGCCCAAAaCGGTTCGGCCATtCTtcGTTACGT
TGGcGCACCCAAAACCAATCCCACGGCCaaCTCAGGcACCCAAaTGCCACCAggcctctc
cacgagtggaactttcaacctctgactgacccagccgccccctggtactcctcatctcgg
agcccccGaTGtCTTGTCgCTGCCTTTCAACATCAGCTTGGCTGGGaCAAGTTTCACTGT
CAATAaCTTTACGTTTGTGCCCCCcaCCGTTCCGgTGTtGCTCCaGATCCTGAGCGGCGC
GCGTaCaGCCcAagagCTCCTTCcGcagggagccgtatacactttgcctccgaataggtc
atcgaaatcactatccctggaggctcaattggaggtccgcatccaatccgccttcacgga
cacgctttctcgg

>contig00518 length=342 numreads=3

gCAGTGgtATCAACGCAGAGTaCGCGGGCAgcTCTGtATATAAcATcGTAGATCCCGTCG
AACGTGaTGTCGTCAACATtGGaGTGgCGGgcGaTAATACGGCAaTTCGATTTGAAaCGG
ACAACTCTGgTCCTTGGATTaCCCATTGAACAATTCTTCACGCACagTGgCcTCGCaGTC
GTAatGGCCGaGGaCaTTAACACGATTGCTACCGAAAaTCCTCCCGCTTAaCGCtCAACG
TTCAGCGTTCAGCttaCAGCCTtCGAATtAGGTCaGATGTGAACAATtAGCtCtgtCCGc
tGTGTtCtCAATTAAAGCAAAACTCCGGAgTCGAGaGttttt

>contig00641 length=604 numreads=8

tgtcaccaaatacatttccttgTTCAGTtCTAAatGaGAAGTtGTTGAGGaTAAtCcAca
CTtAAaTTtGCcTGATTAGACGAAGAAGACTGATCAATCAAAaGTTtGTGGAGGTAAGGC
ATCATAGATAGGgCATAGGTTATCCcAAGCGGTTGGAGGATCCATCGTAGCAATCGTGTT
GACGTCTTCGGCCATGACGACTGCAAGGCCGATCTCAAGATGCCagTCGATATGGCAATG
CATGATCCATGGACCAGAATTATCTGTCTTGAAGCGAATCGTCGTGTTATCGCCAGCCAC
ACCAAtACTGACGACATCACGCTgAACGGGGTTTGCGTAATTATACACAGAGCTTCCTGC
ACTTCgCACAACCGAGAATGCATGGCCGTGCAAATGAATAGGATGAGGACTGCCCACCGC
CCCTCCGGGCATTGAGATCTCGACAACTTTATTAGGTGGCAAGACGTAGACCGAGCCGGC
AGGAAGCAAAGACGCCGCGGTCATTGCACCACTCAtGATTtGTAGAaGAaCAGGTGCAGT
tGGAGGAaCAAatGTGGCGCcGTTGACcGTGAAcTTtAGATCAGTAAaGCTGAACTGGAT
GTCC

>contig00968 length=534 numreads=6

tCCCCgCAGGATCGATCTATGAGTTGCCACCGAACTCCGTTATCGAATTATCTATGCCAG
GTGGATCGGCAGGTAGCCCTCATCCAATCCATCTTCATGGGCACACCTTCTCTGTCATAC
GAAGCGCGGGAAGCAaGATATACAACTACAACAACCCCGTCCGTCGGGACGTCGTTAGCA
TTGGTGATTCTAACGATAATGTTACGATCCGGTTcGAAACCGATAACGCTGGGCCTTGGC
TTCTCCACTGTCATATCGACTGGCACCTTGATCGAGGTCTTGaTGTTGTCTTcGCTGAGG
ATATCCCGGACATTGCCAAGCAGAaTCcTCCcATTGCTTGGGAGAATCTCTGCaGGgaCT

AaCATGTCGATGGGAACCCAAGAAAGTAATCAACTCACTAGTCCTGGGCATTGATTTaaG
aCGTTGATTGATCCGGCCATGACTTCGTACAAtGTTATTCTGGGtATTTGTtATTtttGT
tATCTATATCGTTAtCgtcacagtattccgagtacgagaggtggatactatttc

>contig01153 length=457 numreads=3
TTGAGTTCCAgAgTtCGGCAAGGCCcGTATCcaGTAGTtGGTGACTGGCTGGTtGGTACt
GAGCaCGAATGAATAGCGTtGGgCCAcgAaGATTtGGATCGAATCGACGGTAAGTGGCTG
AACGTTCACGCCATCAACTTCAATGATAGTCATGGTgTGGCCATCGATGGAAAAcaCAtA
GTTGGGAtCGCAGGAGAGGGAAACAAGACGGAAgCGATACcGCTTATTGGGAAgAACGGT
AATCACAGAAAGCGGCGATGTtGGgCCACCAGCATAGCGGCCTTTGCCGTTAATCAAGGT
TGCGTCTGCcGTtGGaaCcAGTCCcGCCGaTGGCGCaGGTgTATGATACCAATCTCcAAG
AGTTATGACAGTGGTGTCATCGTCGAAaTCAtATTtGGATCGGTGAGGGTCTTTTCGGTC
ATAGATGaCcAAAgCACCTCGGAGGCCgTCAcaGTAt

>contig01577 length=321 numreads=2
ATAaGGTCATCGAAATaTCGAtGCCaGGAGGGTCAACTGGCAGTCCCCTCCTATCCATCT
TCACGGCCACACATTCTCTGTCATCaGAAGCGCGGGGAGCACTGAATACAACTACGtCAA
CCCAGTTCAGCGTGATACTGTTAGCATAGGATTAGCTGGCGATAACaCCACcATCCGCTT
CAAGACTGATAATGCCGGTCCTTGGATCATGCaCTGTCaCATCGACTGGCATCTCAaTaT
TGGTcTCGCcGTCGTCaTGgCCGAaGATGTTGCAaCGATTGCCACaGCGAATCCTCCCAC
TGCTtGGGaTAATCTGTGCCc

>contig03441 length=800 numreads=11
CAGCAGaCTCAGTGTCATTTCCGGGACCACTTATCGTCGGgAAaCACTGGCGATGTCTTC
ACCATCAACGTTATTGATGAATTGGTCGATCAGACTATGCTGACGAGTACTAGCATTCAT
TGGCATGGCCTTTTCCAGAAGGGCTCAGATTGGGCTGAtGGgCCAGTCGGTGTAACTCAG
TGCCCTATCAAACCaGGAAATTCTTTTCAATATCAATTCTCTGTCCCaGATCAAGCTGGA
ACATTTTGGTACCACTCGCACGACTCTACTCAATACTGTGATGGCCTAAGaGGGGCgCTC
GTGGTGTaCGATCCGAACGATCCTTACAAACCTCAGTACGACTTTGACGATGACTCTACC
GTCATCACTCTCGCTGATTGGTATCATGCACCTGCTACCCTATTGGGAATGACCCCAAGA
GCGCCCACACCACTGGCCACCTTGATCAATGGCAAAGGACGTTATGCAGAGGACCCAACT
GCTGAGCTCGCAGTCATATCAGTCATACCAAACAAGCGTTATAGATTCCGCTTAGTGTCt
ATTTCATGCGATCCAAACTTTGAATTCACGATCGACAATCATAGCATGATGATTATCGAA
GTCGACGGaCAGAATATCCAACCTCTTaCtGTTGACTCAATCACTatATCTGCAGGACAG
CGCTACTCCTTTAtCCTCCAGGCTAATAAaCAGGTGTCCAACTACTGGATTCGCTCGCTC
CCcAacAGTGGCCcACCAGGGTTCACCAaCGgCGTCAATtCTGCAAtCCTTCGCTACGTT
GGAGCGCCCATCGCCGAccc

>contig04378 length=627 numreads=10

CCCATCTTCTGGTCGCTCAACTGAATCATGCTTTCCCTCTCAGCCATTGTTGGCCTTTCT

GtCGTtGTTGGAGCCTATGCTGCCATCGGTCCCACTGCCAATGTTTAtATCAaGAaTGCT

GCTATCGCGCCCGATGGATTCACTCGCTCcGCCGTATTGGCGGgAGCATCCTCATCGaGT

GTAAaGTTtCCTGGCCCACTTGTCACTGGCTTCAAGGGaGATACATTTAGCTTGAATGTt

ATTGACCAGCTCACCGATAACACCATGTTGCTGAGCACAAGTATTCATTGGCACGGTaTG

TTCCAAGAAGGGACTACATGGgCCGATGgTCCAGTGGGCGTCAACCAGTGCCCtATCGCC

CcGGgCCACTCCTTCTTGTACAaGTTtAGCGTCCcAGATCAaGCGGGgACGttctGGTAT

CAcTCCCATTTTtCAACGCAGTATTGCGATGGCCTTCGAGGCGCCATGGTTGTCTACGAC

CGCCATGATCCTCATCGCtCTAGaTATGATTTtGATGACGAGAGCACTGTCATCACCcTG

GCTGATTGGTACCatACTGCTGcACCcATTgCGGgCcTTGTGCCTACGGCAGATGCTACA

CtGATcAATGGCAAAGGCCGCTaTGCt

>contig01192 length=604 numreads=5

gggCATCCTGGgAGTGGgTGGCAAGTGGCCAGAGTACGTAGCGCCGAAGGAAGGCGATAG

tCGATGCTCGTGTCCCGCGCtGAACGCGATGGCTAACCATGgTATCCTCCCTCACGATGG

CAAAAACaTTAAATtCAATGAGTTCGgCACCAAAatCAAGGgCACATACAATTTCGCATC

CAGTTTCTGTTTCTtCGtCCCCAACTTCGCCgCCAACATGTtGAAGCGCAGCtACAGCAA

AGACaTCTTCGACCTCGCGGACCTTGATCTCCACAATGGCATAGAACACGACGCATCTCT

TACTcGCGAAGATAGTGCCCTTGTACCCAGTCAAGCAAAGCCCcACGTTCCTTTtATTAA

AGAGCTTCTGGATTGTGCTTCTGGAAAGGaCAAGGATGGCAATGTTATCATTACGGCTAa

GGATCtAtCAGACTACTCTGCGAAGCGCCGCGTCGaCGCTCGCGCCaCAAaCTCTGATTA

CACAcTAGacacaTTCCACAAAGTATtCGGCAGTTCCAaTTCCTCGACCCTCcTCACGAt

CTTCGGAGGCcGTGtGGaCGATTTGGAAACGTTCTTGATCGAGgAGcGCATTCCTGATGG

gTgg

Figure 6.6 Primer sites for all selected contigs of *S. aeruginosa* cDNA sequencing.

Scheme presents primer sites for inverse PCR for each contig. Forward primers are indicated in red, reverse primers are marked in blue. Uppercase letters in sequence indicates more "certain" regions of the contig, lowercase letters indicate less accurate regions of each contig.

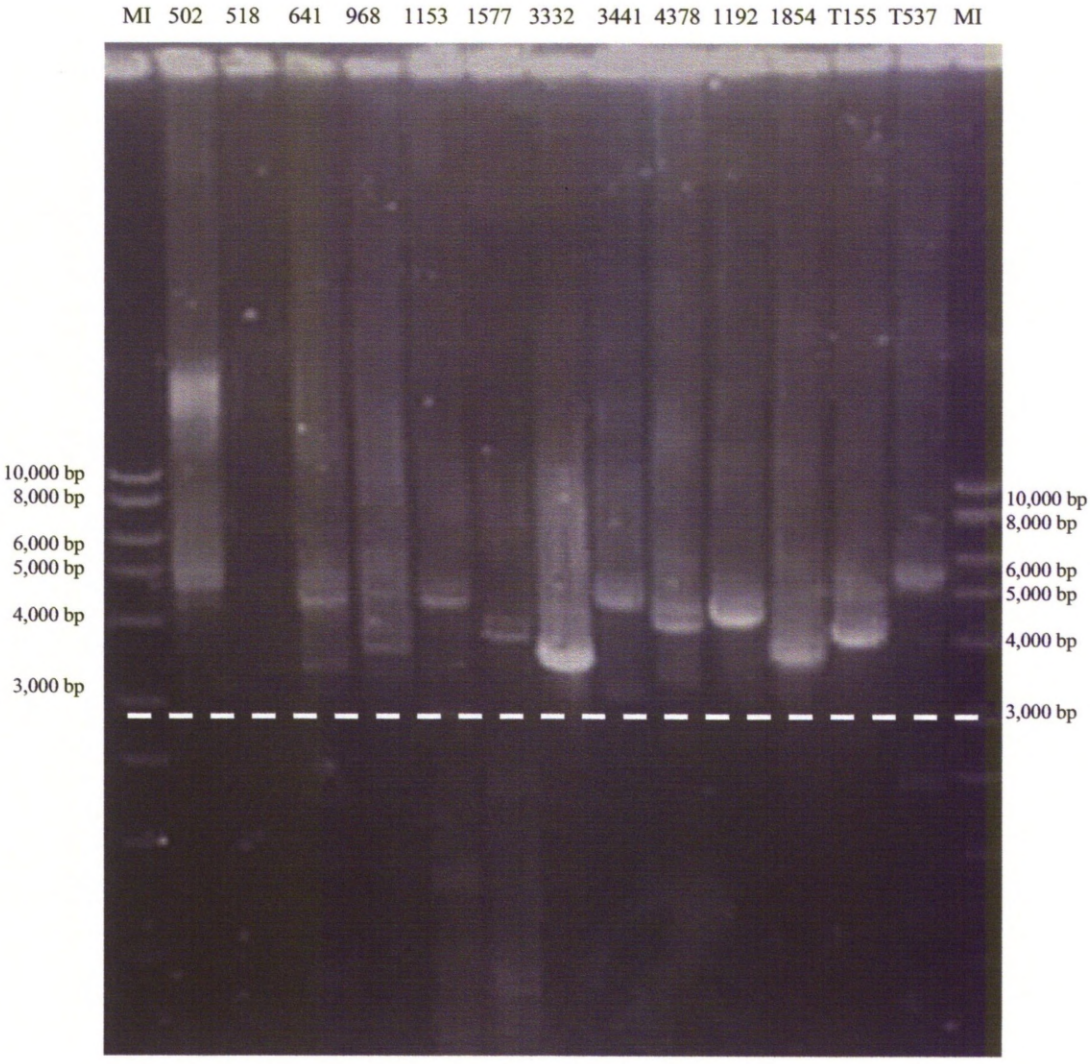


Figure 6.7 The result of inverse touchdown PCR amplification.

MI – Hyperladder I (range 10000 – 200 bp); 502 → T537 – inverse touchdown PCR products from their respective contigs; 502-4378 - *S. aeruginosa* laccase contigs; 1192 and 1854-*S. aeruginosa* putative haloperoxidase contigs; T155 and T 537 *Geomyces sp. P7* contigs. Sample 518 failed to amplify with touchdown PCR, amplification at the annealing temperature 62 °C was successful. Samples 3332 and 1854 were false annotation positives; T155 and T537 were the amplifications of *Geomyces sp. P7* esterase and lipase respectively. The dashed white cut-off line shows the size of an empty pBluescriptSK(-) vector. All PCR products considered positive must have the size larger than 3 kb. Separation was performed on 1 % (w/v) agarose gel.

7. Assembly of the genes.

After recircularization, all plasmids were transformed into competent cells, isolated and sequenced. Unfortunately most of PCR products were not specific enough yielding incomplete gene sequences, despite this approach working very well for the aforementioned gene of the lipase from *Geomyces sp. P7*. Because of this problem, assembling known contigs with the sequencing data from the respective plasmid were not as straightforward as expected. It is thought that this problem probably occurred due to the existence of multiple genes of very similar laccases and/or multiple splice variants of these laccases, which may have caused these problems; however there is no clear evidence that can support this reasoning, but no problems of this kind were encountered when generating full length sequences for *Geomyces sp.P7* genes which were very straightforward, no multiple gene variants were detected for this project.

To address this problem, all data collected from plasmid sequencing were divided into groups representing 5′ and 3′ fragments of the gene and added to the database of already selected contigs. All these data were assembled *in silico* using the CAP contig assembler from Bioedit program. When initial assembly was completed the singleton database was searched for additional sequence data that could expand the sequence data. As a result of these attempts two full length sequences of laccases were assembled *in silico* along with an array of incomplete sequences missing either the 5′ or 3′ end of the gene. Longer sequences composed of these data were denoted as maxicontigs (for large incomplete sequences) and supercontigs (for complete sequences). Figure 6.8 presents typical assembly of a gene using data from contigs, singletons and plasmid sequencing. Full assembly can be found in Appendix 1.

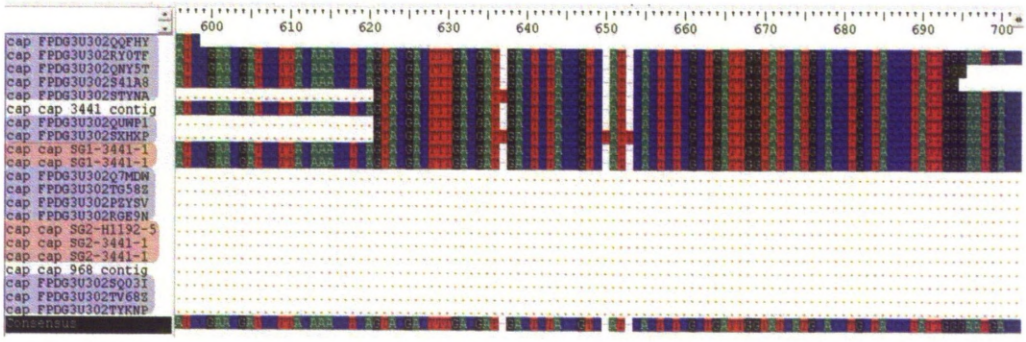


Figure 6.8 The typical assembly of a supercontig with CAP contig assembler module of Bioedit.

Singleton sequences are highlightd in blue; plasmid sequences in red, original contigs assembled by Dr Kevin Ashelford remain white.

The final result of the assembly is summarized in Table 6.4. The table presents the length of maxicontigs and supercontigs and representation of which part of the gene according to homology searches they are. The annotation represents the respective primer name used for the amplification i.e. N1C2 was amplified with primers LacNterm1_F; LacCterm2_R etc. The primer sequences can be found in M&M 3.23.

Table 6.4 Summary of data gathered from laccase sequencing assembly

Annotation	Length [bp]	5´ ATG	Middle	3´ Stop	Amplified
N1C2	1539	+	+	+	+
N2C1	1554	+	+	+	+
NMC1	1317	-	+	+	+
NMC2	858	-	+	-	+
NMC3	1320	-	+	+	+
NMC4	1293	-	+	+	+

+ – the feature is present in maxicontig or supercontig
- – the feature is present in maxicontig and supercontig
N1, N2 – known N termini of translated proteins (BLASTX analysis)
C1, C2 – known C termini of translated proteins (BLASTX analysis)
NM, CM – unknown N/C termini, known internal sequence of translated proteins (BLASTX analysis)
Amplified – a sequence was successfully amplified with given primers

To obtain full length sequences of the incomplete laccases several methods to amplify missing fragments were attempted. Unfortunately none of these attempts succeeded. To address the problem a probe was designed to screen the cDNA library of colonies. The probe was designed to amplify the most homologous region of all *S. aeruginosa* laccases. To do so all assembled sequences were translated to protein sequence and multiply aligned, the fragments showing highest sequence similarities were chosen as sites for PCR amplification. Figures 6.9 and 6.10 present multiple sequence alignment and fragments chosen for primer design. Figure 6.11 presents attempts to amplify full length sequences from cDNA libraries, two full length laccase genes are presented along with the middle fragment amplified with LacMid1 primers

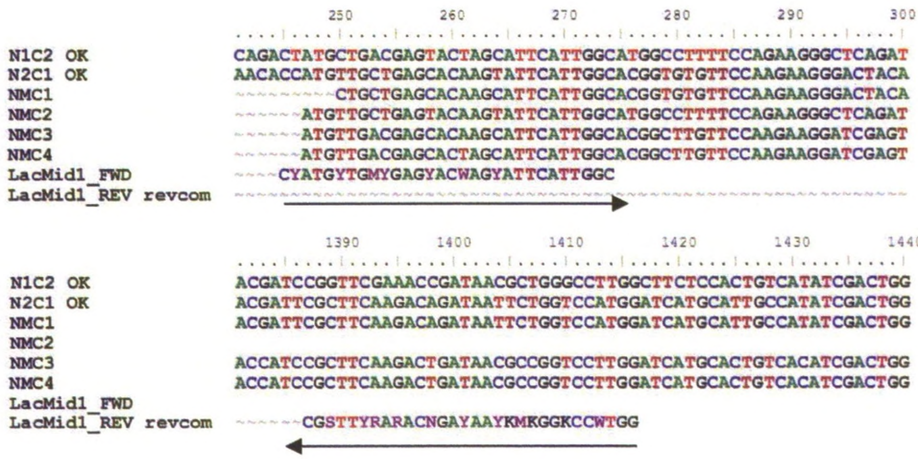


Figure 6.9 Multiple sequence alignment of DNA fragments of contigs used for primer design to amplify middle part of all laccases.

LacMid1_Fwd - forward primer for the amplification; LacMid1_Rev revcom – reverse primer presented in the reverse complement form that shows sequence similarity between the primer and laccase sequences. Arrows show primer orientation.

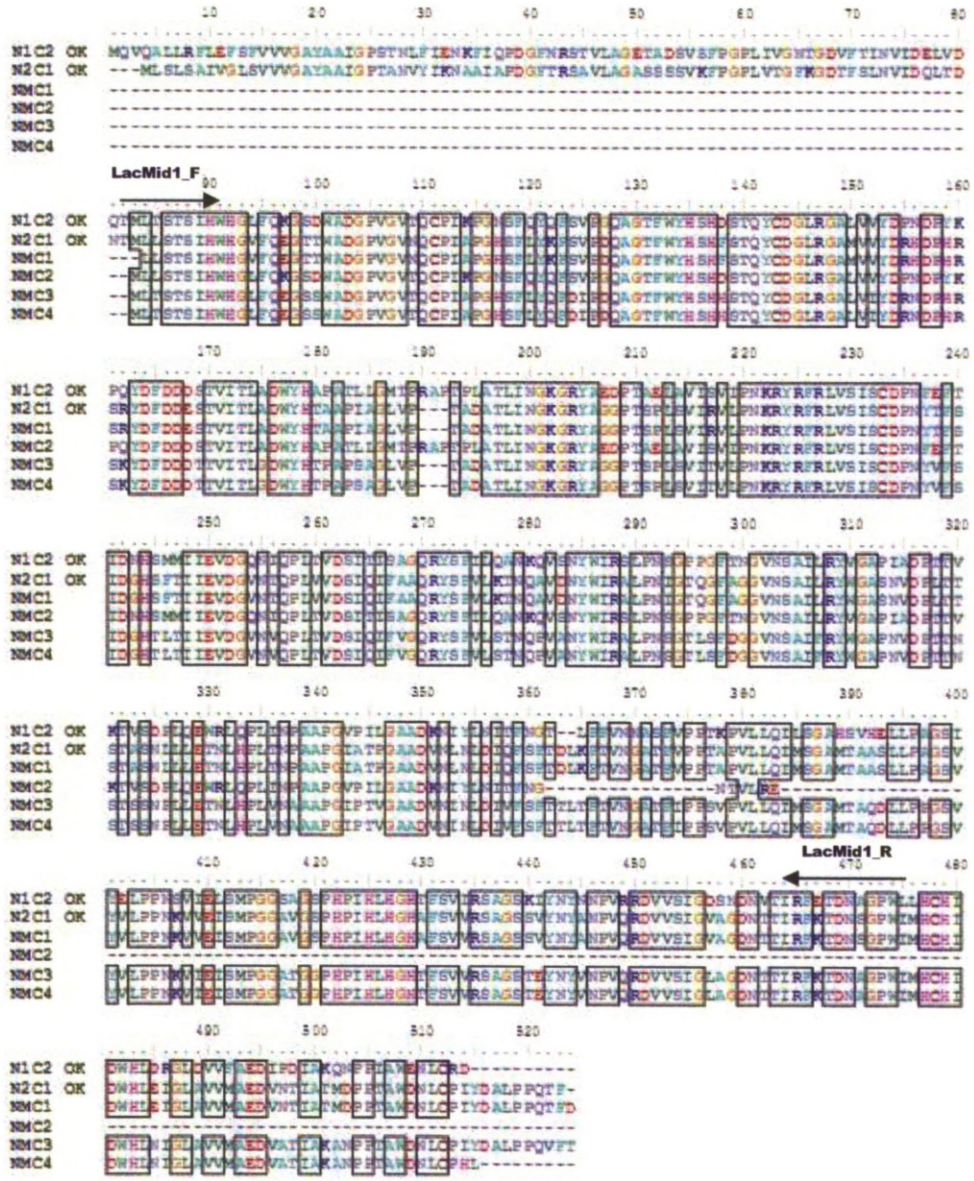


Figure 6.10 Multiple sequence alignment of assembled laccases.

Contigs were assembled and amplified with respective primers. Sequences sharing over 80 % of sequence similarity are marked in boxes. Protein fragments selected for primer design to amplify the middle part of laccases, presented in Figure 6.9 are marked with arrows, names of primers presented above.

Figure 6.11 Amplification products.

MI – Hyperladder I (range 10000 – 200 bp); N2C1 – amplification of *S. aeruginosa* N2C1 gene; N1C2 – amplification of *S. aeruginosa* N1C2 gene; NMCM – amplification of *S. aeruginosa* homologous middle fragment with degenerate primers LacMid1. Separation was performed on 1 % (w/v) agarose gel.

Hybridisation experiments using the NMCM probe to screen the plasmid cDNA library that were supposed to provide us with the missing genetic information regarding the genes for the laccases did not show any success despite several trials; furthermore the probe highlighted a few false positives that did not show any sequence similarity to laccases after sequencing. This was probably caused by poor probe choice and the size of the library. Based on these results a shorter probe of approximately 100-200 bp and a partial cDNA library limited to DNA size of around 1500 bp to limit the complexity of tested cDNA is suggested. This subtracted

cDNA pool could be obtained either by size exclusion chromatography or electrophoresis coupled with band excision. The subtracted cDNA pool could be then ligated to plasmid vector, transformed, plated and tested *via* colony hybridization. Unfortunately it was impossible to perform this experiment due to lack of time.

8. The comparison of genetic and proteomic data.

When all the genetic information available was collected together it was decided to compare the sequences with proteomic data gathered earlier. To do so the translated protein sequences were combined into a mini-database and a BLASTP search was performed using the Bioedit program with a short sequence specific PAM 40 matrix. Sequence similarity between peptides obtained through proteomic analysis and summarised in Chapter 4 (Figure 4.30 and 4.31) was analyzed. A flowchart presenting the summary of all results collected for *S. aeruginosa* laccases is presented in Figure 6.12. Top part of the flowchart summarises the biochemical properties of these proteins. Bottom part of the flowchart shows sequence data generated at protein and DNA level. Thick dashed lines show identity at the protein level between different methods of analysis. The peptide GPSTNLFIENK (previously annotated as GPNANLFLQNK – sequencing spectrum was inconclusive) is an essential one for this dataset. It was the only one of the good quality found when purified proteins were analysed by tandem mass spectrometry. It was also found in proteomic analysis of isoforms on 2D gels. A full length sequence of a laccase N1C2 also contains this sequence at its N terminus. Other data are partial, a peptide LTVKVPLRT was found in both glycosylated and non-glycosylated fractions analysed with 2D gels. Two other peptides (TEYNVNV and TSPLSVLTVL) found with this analysis were found in partial DNA sequences assembled from contigs and inverse PCR sequences. Thin dashed lines show identity between protein sequence and

translated DNA sequence. In order to have a full picture of laccases produced by *S. aeruginosa* it will be essential to obtain the missing data.

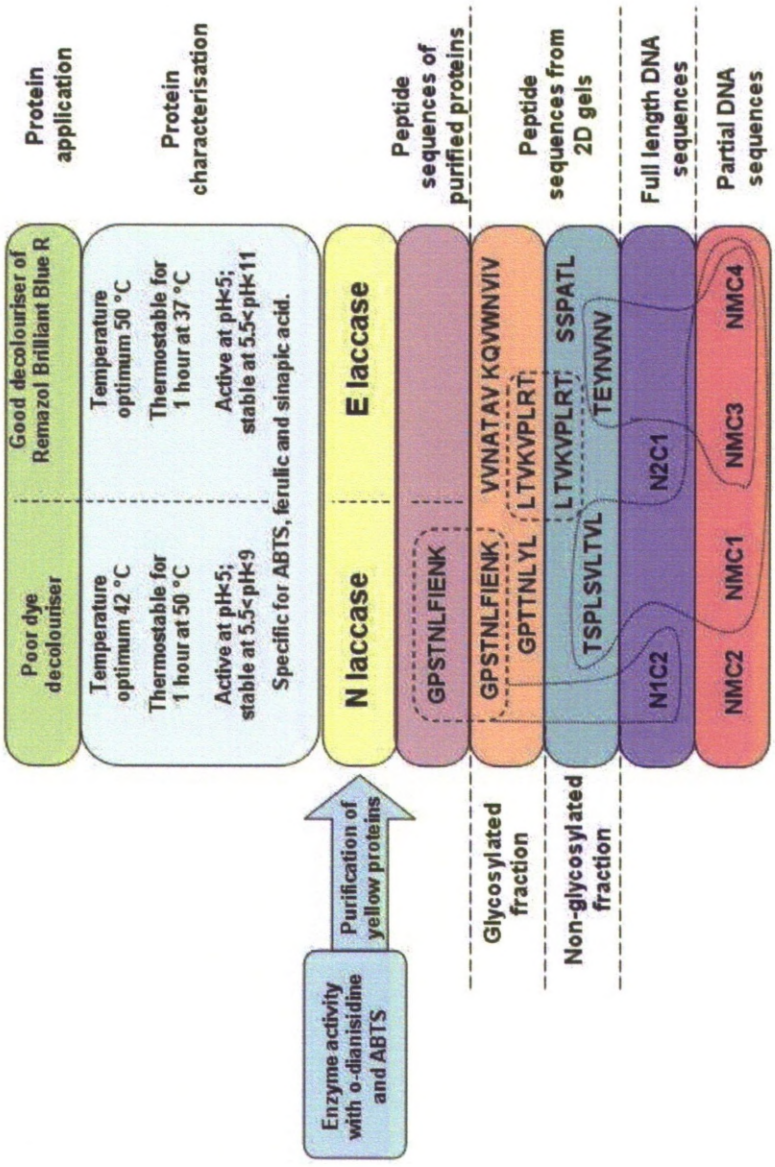


Figure 6.12 Summary of all findings about *S. aeruginosa* laccases.

Thick dashed lines show identity at the protein level, thin dashed lines show identity between peptide sequence and a translated DNA sequence. Types of experiments are shown in bold at the right side of the scheme.

9. The putative haloperoxidase of *S. aeruginosa.*

During the sequence analysis of *S. aeruginosa* a possible haloperoxidase that might be the protein first reported by Moore (Moore, 2007) was found. A gene similar to the *C. fumago* chloroperoxidase was found on contig 1192. When assembling the genes of laccases using CAP contig assembler of Bioedit program the identical procedure was performed for contig 1192. As a result a longer DNA sequence was obtained containing the full 3′ end of the gene including the stop codon and an in frame methionine that could be the start codon, although the homology search was inconclusive as to whether this first methionine could be the first amino acid of the protein. The assembled sequence was compared with *C. fumago* chloroperoxidase and two other putative haloperoxidases from *Agaricus bisporus* and *Aspergillus nidulans*. The multiple sequence alignment of these proteins is presented in Figure 6.13. Several primer pairs were designed to amplify this gene; however the amplification process was not successful and resulted in truncated sequence. An alternative method e.g. colony screening with long PCR amplified probe is recommended.

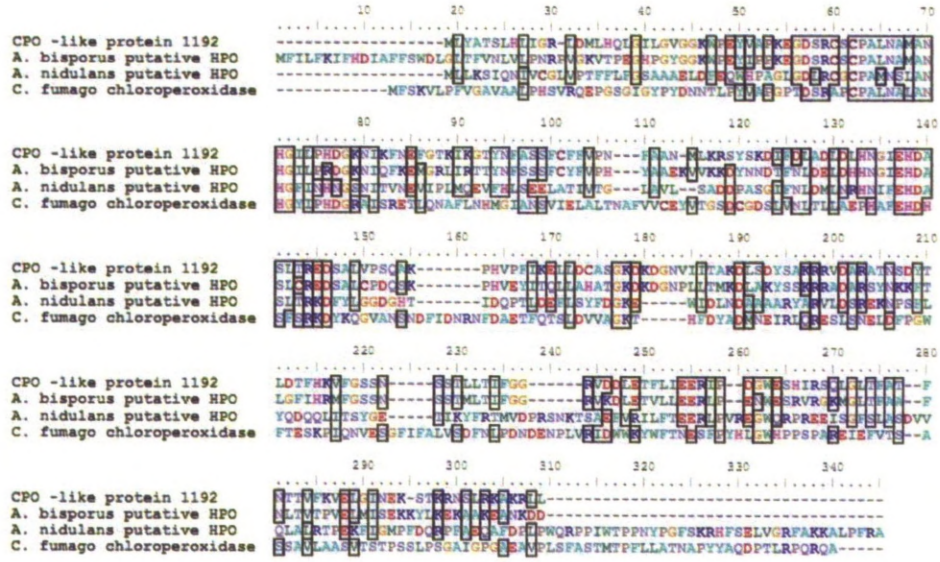


Figure 6.13 Putative fungal haloperoxidases.

Multiple sequence alignment of *in silico* assembled *S. aeruginosa* CPO-like protein with *C. fumago* haloperoxidase, *A. bisporus* putative haloperoxidase a gene of which Moore was trying to amplify and *A. nidulans* AN 7823 protein showing homology to CPO.

10. Summary.

Unsuccessful attempts to find the genes encoding the laccases of *S. aeruginosa* and the arrival of Mr Tomasz Florczak with a new project focused on finding the gene coding an unusual thermostable lipase from *Geomyces sp. P7* prompted us to abandon the traditional methods of isolating the genes based on protein sequence, library screening and homology. To obtain rapid results at a reasonable cost we decided to explore new avenues opened by the introduction of 454 pyrosequencing. The cDNA of the organism was sequenced in order to find fragments of genes with homology to known laccases and lipases. To identify the cDNA fragments a BLASTX analysis of all contigous sequence fragments (contigs) was performed. The analysis of *S. aeruginosa* sequences resulted in identification of eight contigs homologous to laccases and one similar to a haloperoxidase. Those selected contigs served as a target sequence to perform inverse touchdown PCR on circular plasmid cDNA libraries. As a result of amplifications more data was gathered. Unfortunately some sequences still remained incomplete. That was probably due to either very high homology between laccases or problems in the PCR. The problem was partially solved by *in silico* gene assembly using the entire pool of 454 sequencing data. Using this approach two full length sequences were obtained and four partial sequences. The comparison of translated protein sequences obtained through this approach with proteomic data obtained earlier proved that one of the full length sequences probably corresponds to the N laccase isolated earlier. Several partial sequences (NMC1, NMC3 and NMC4) matching proteomic data were isolated. Isolation of full length sequences of these proteins is necessary to confirm these findings To obtain more data a strategy to colony screen the library with a complex probe created by PCR amplification of all homologous sequences was devised. Unfortunately this attempt gave many false positives, because of probe length and library complexity. However if shorter probe is designed

and applied on a subtractive cDNA library it should be successful A similar approach should be employed for the putative haloperoxidase.

Chapter 7

THE BIOINFORMATIC ANALYSIS OF *S. AERUGINOSA* LACCASES

1. Introduction.

The successful isolation of two genes encoding laccases of *S. aeruginosa* allowed a variety of bioinformatic analyses to be performed to further our knowledge of these proteins. The first analysis focused on identifying intronic sequences within both genes *via* aligning gene sequences with a PCR product amplified with the same primers from genomic DNA. Then an attention was focused on protein sequence and the identification of potential glycosylation sites and copper binding sites using multiple sequence alignments and sequence analysis tools available online. Finally the FRankenstein Monster's approach was applied to model the molecular structure of both proteins in the hope of unveiling the structural factors that might contribute to non-blue character of this protein. This method selects modelling templates on the basis of fold recognition rather than sequence similarity.

2. Finding intronic sequences of *S. aeruginosa* laccases.

Genes of ligninolytic fungal enzymes are well characterised for their multiple splice variants and the number of introns per single gene. To find the number of introns in the two full length genes isolated, the genomic DNA of *S. aeruginosa* was extracted; both genes were amplified from this template using (LacNterm1_FWD, LacCterm2_REV and

LacNterm2_FWD LacCterm1_REV), cloned into pJET 1.2 and sequenced. The sequences were then compared with gene sequences amplified from cDNA and aligned with each other to identify introns. As a result it was determined that the gene of laccase N1C2 contained 15 exons and 14 introns, gene of laccase N2C1 contained one less with 14 exons and 13 introns. The shortest exon was just 15 nt in case of N1C2 and 17 nt for N2C1, the longest exons were 526 nt and 524 nt respectively. The majority of laccase introns follow the general GT-AG rule that describes sequences at both ends of an intron. However 5 intronic sequences of N1C2 and 2 N2C1 introns differ from this rule without showing any other pattern Figure 7.1 presents these findings for first 350 bp of sequence. Entire sequence can be found in Appendix 1.

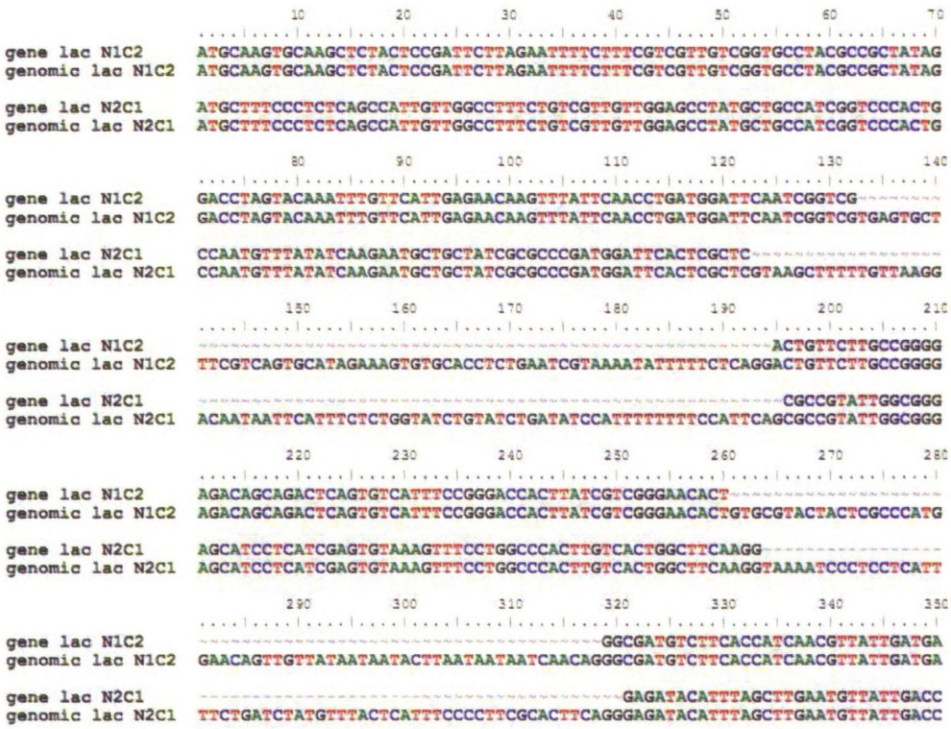


Figure 7.1 Exons and introns of *S. aeruginosa* laccases.

Genomic sequences of *S. aeruginosa* laccases show multiple introns that partially follow GT-AG rule.

3. Protein sequence analysis of *S. aeruginosa* laccases.

The sequences of both laccase genes were translated into proteins and analysed *via* variety of tools available: Signal IP 3.0 to detect signal sequences for secretion (Bendtsen et al., 2004), Protparam to calculate basic parameters of the protein product (Gasteiger et al., 2005) and NetNGlyc to find putative N glycosylation sites (Gupta et al., 2004). The first feature analysed was the presence and length of signal peptides. Both protein products have relatively short signal peptides of 21 and 18 amino acids and both posses an identical protease cleavage site AYA-AI (where "-" is the cleavage site) This sequence follows the general pattern (called -3, -1 rule) for signal peptide cleavage sites that require small neutral amino acids (like alanine) at sites -3 and -1 with respect to the cleavage site. The sequence at the ends of the cleavage site should be composed of a positively charged N-terminal region, followed by a hydrophobic H (helical) region and neutral but polar C-terminal region (Nielsen et al., 1997)., what is generally true for *S. aeruginosa* laccases.

The basic biochemical data such as molecular weight and theoretical pI were estimated as 53.7 kDa and 5.1 for the protein product of N1C2, those parameters of N2C1 were calculated to be 53.6 kDa and 5.7 respectively. The molecular weight of the proteins is in agreement with biochemical characterisation of the purified proteins, whereas the pI calculations are around 1.5-2 pH units higher then expected. This is probably due to the fact that whereas theoretical calculations are performed as a summary of all charges of individual amino acids, real life experiments were performed on native proteins, therefore the surface charge is likely to have had an effect on the migration rather than the net charge.

N-glycosylation analysis showed that N1C2 possesses four putative N-glycosylation sites whereas N2C1 is not glycosylated. Figure 7.2 summarizes these findings in a graphical manner.

N1C2

MQVQALLRFLEFSFVVVGAYAAIGPSTNLFIENKFIQPDGFNRSTVLAG
ETADSVSFPGPLIVGNTGDVFTINVIDELVDQTMLTSTSIHWHGLFQKG
SDWADGPVGVTQCPIKPGNSFQYQFSVPGQAGTFWYHSHDSTQYCDGLR
GALVVYDPNDPYKPQYDFDDDSTVITLADWYHAPATLLGMTPRAPTPLA
TLINGKGRYAEDPTAELAVISVIPNKRYRFRLVSISCDPNFEFTIDNHS
MMIIEVDGQNIQPLTVDSITISAGQRYSFILQANKQVSNYWIRSLPNSG
PPGFTNGVNSAILRYVGAPIADPTTVKTVSDPLQEWRLQPLINPAAPGV
PILGAADKNIYLNITFNGTLFFVNNASFVPPTKPVLLQILSGAHSVHEL
LPAGSIYELPPNSVIELSMPGGSAGSPHPIHLHGHTFSVIRSAGSKIYN
YNNPVRRDVVSIGDSNDNVTIRFETDNAGPWLLHCHIDWHLDRGLDVVF
AEDIPDIAKQNPPIAWENLCRD*

N2C1

MLSLSAIVGLSVVVGAYAAIGPTANVYIKNAAIAPDGFTRSAVLAGASS
SSVKFPGPLVTGFKGDTFSLNVIDQLTDNTMLLSTSIHWHGVFQEGTTW
ADGPVGVNQCPIAPGHSFLYKFSVPDQAGTFWYHSHFSTQYCDGLRGAM
VVYDRHDPHRSRYDFDDESTVITLADWYHTAAPIAGLVPTADATLINGK
GRYAGGPTSPLSVIRVLPNKRYRFRLVSISCDPNYTFSIDGHSFTIIEV
DGVNTQPLVVDSIQIFAAQRYSFVLKTNQAVDNYWIRALPNIGTQGFAG
GVNSAILRYWGASNVDPLTTSTASNLLLETNLHPLTNPAAPGIATPGAA
DVNLNLDIQFSFTDLKFTVNGATFVPPTAPVLLQIMSGAMTAASLLPAG
SVYVLPPNKVVEISMPGGAVGSPHPIHLHGHAFSVVRSAGSSVYNYANP
VQRDVVSIGVAGDNTTIRFKTDNSGPWIMHCHIDWHLEIGLAVVMAEDV
NTIATMDPPTAWDNLCPIYDALPPQTF*

Figure 7.2 Summary of protein sequence analysis findings.

The figure presents a summary of the bioinformatic findings on the two *S. aeruginosa* laccases N1C2 and N2C. Key : Signal peptide, signal peptide cleavage site AYA/AI, N-glycosylation sites.

4. Multiple sequence alignments of *S. aeruginosa* laccases.

To compare the laccases of *S. aeruginosa* with enzymes from other organisms and allow molecular structure modelling multiple sequence alignments of protein sequences were performed. The analyses were performed with ClustalW module of Bioedit. Alignment was made on the basis of BLASTP results and refined to proteins about which some detailed information was available (publications, direct sequencing submissions, etc.), excluding hypothetical proteins from genome and metagenome sequencing projects. The most important fragments of these alignments showing amino acids coordinating catalytic coppers are presented in Figure 7.3; the full alignment is presented in Appendix 1. Both multiple sequence alignments were combined to create a phylogram with ClustalW2 from EMBL website. The phylogenetic tree of these proteins is presented in Figure 7.4. Phylogenetic analysis suggests that even the best modelling templates are quite distant from the laccases of *S. aeruginosa* as they occupy relatively remote branches from all known crystallographic structures.

The multiple sequence alignment presented in the Figure 7.3 shows the most conserved region of the protein around the T2/T3 copper binging histidines of WYHSH motif. The analysis of the alignment indicates that at the sequence level the proteins of *S. aeruginosa* do not seem to differ very much from laccases from other ligninolytic fungi e.g. *Trametes versicolor*, *Pleurotus ostreatus*. The alignment proved that all the essential amino acids of *S. aeruginosa* laccases associated with copper coordination are conserved with respect to other protein sources. This is summarized in Table 7.1. Four highly conserved motifs are responsible for metal coordination in laccases. The sequence of those motifs remain conserved in laccases of *S. aeruginosa.*

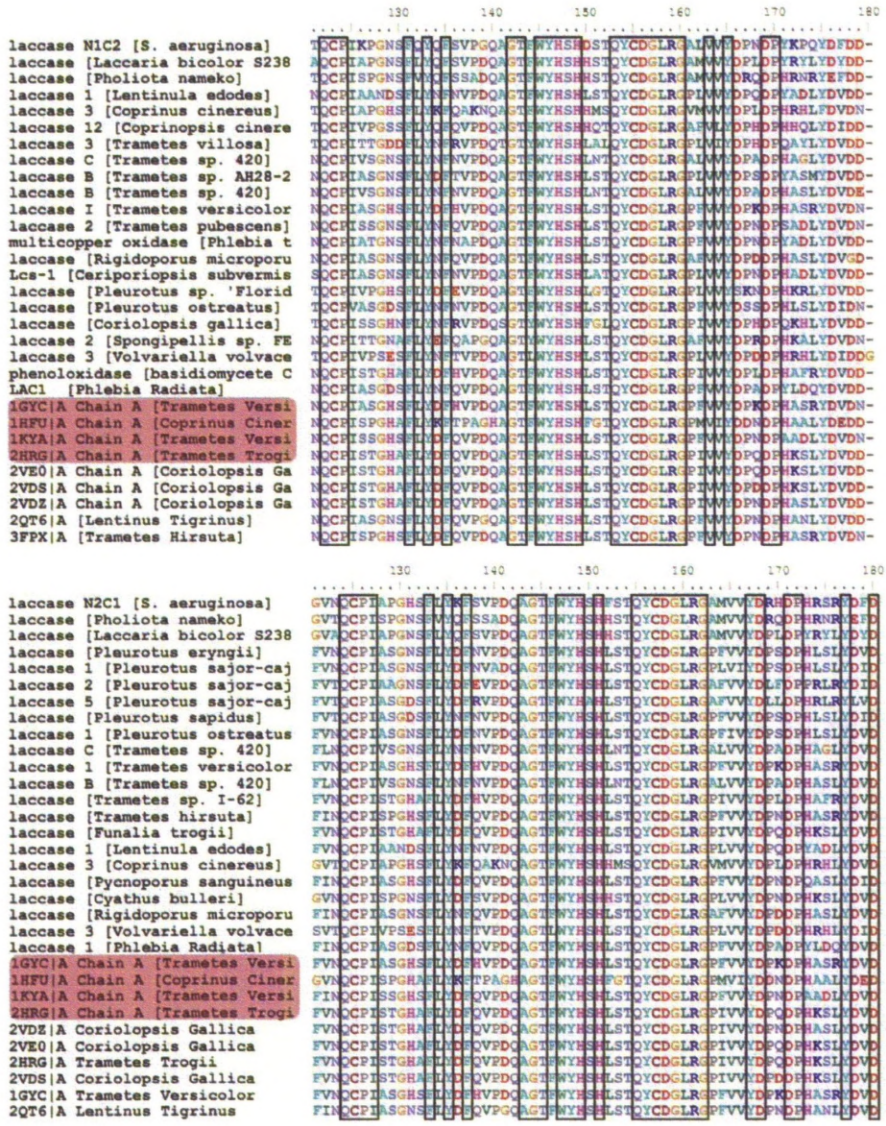


Figure 7.3 Multiple sequence alignment of *S. aeruginosa* laccases (N1C2 on top, N2C1 at the bottom) and the most homologous proteins.

The protein sequences for the alignment were selected on the basis of the BLASTP score and represent the most sequence similar proteins. The figure presents the copper T2/T3 environment. Identical amino acids are marked with boxes. The crystal structure templates selected for the molecular modelling step are marked in red.

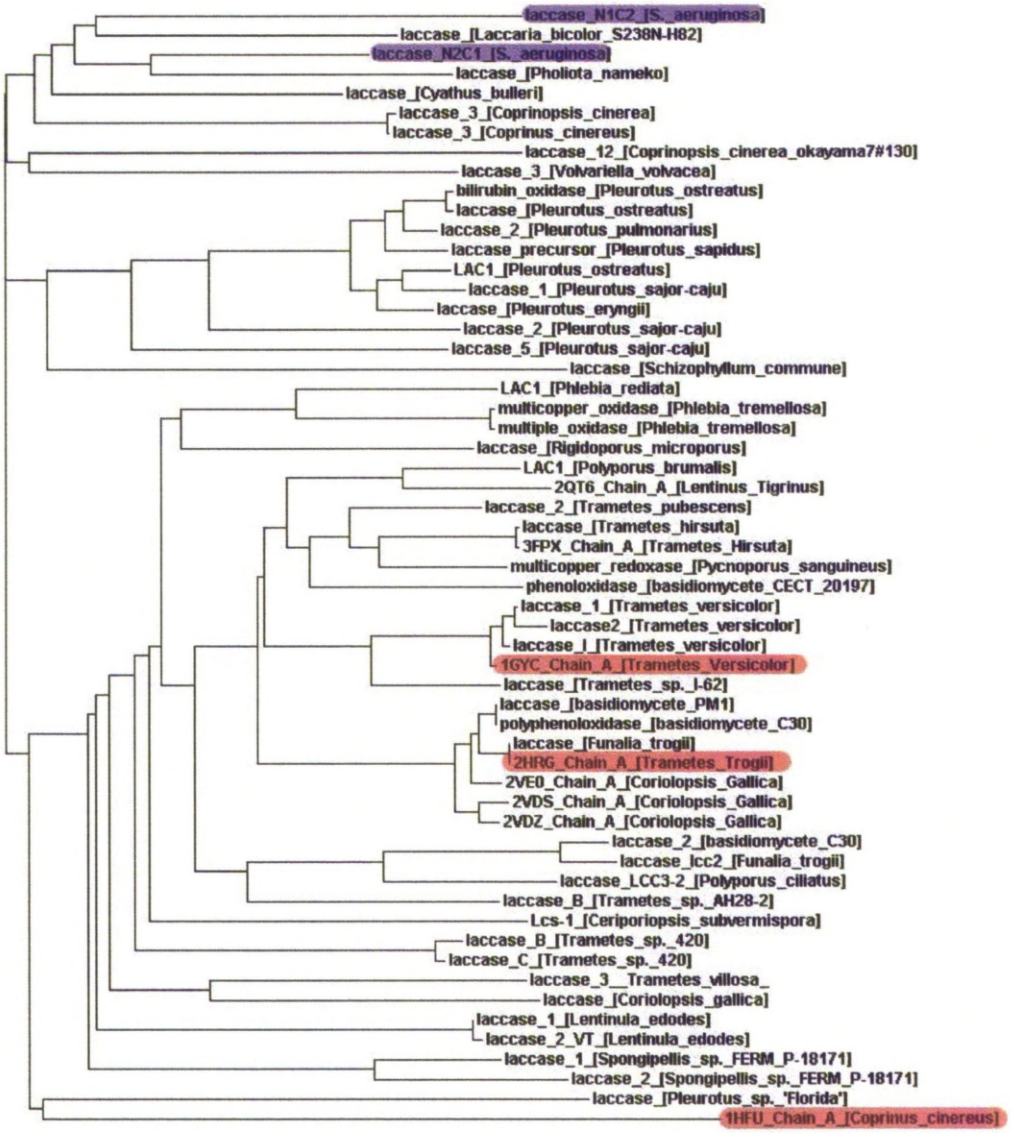


Figure 7.4 Phylogram of laccases used to create multiple sequence alignments.

Laccases of *S. aeruginosa* are highlighted in blue and modelling templates in red. Phylogram shows that laccases of *S. aeruginosa* are quite distant (within the laccase family) from all available modelling templates.

Table 7.1 Copper coordination in *S. aeruginosa* laccases

Amino acid sequence	**Position of first amino acid***		**Function**
	N1C2	N2C1	
PHPIHLHGH	398	397	T1 copper coordination
HCHIDWH	454	453	T1 copper coordination
TSIHWHG	66	66	T2 copper coordination
PHPIHLHGH	398	397	T2 copper coordination
TSIHWHG	66	66	T3 copper coordination
HCHIDWH	454	453	T3 copper coordination
GTFWYHSH	109	109	T3 copper coordination
PHPIHLHGH	398	397	T3 copper coordination

[*] aminoacids numbered as in mature protein without signal peptide; numbering starts with first letter of motif

Conserved residues are underlined; residues that correspond to the function indicated in the column are marked in **grey**

Having confirmed that the copper environment (at least at the sequence similarity level) is very similar to other laccases, further sequence analysis was performed, focusing on similarities and differences of *S. aeruginosa* proteins from other fungal laccases. Several regions of high sequence similiarity were found, however a literature search did not provide any valuable information about the function of these regions. It was decided to review their function when molecular models of *S. aeruginosa* proteins were completed and compared to known 3D structures of laccases.

5. Comparative modelling of *S. aeruginosa* laccases.

5.1. Model building.

In order to obtain more detailed information about the laccases of *S. aeruginosa* it was attempted to build molecular models of these proteins using the FRankenstein Monster's approach (Kosinski et al., 2003). The first step involved selection of templates for modelling using fold recognition techniques provided *via* the Phyre server (Kelley and Sternberg, 2009) that collects three methods of fold recognition: Psi-Pred, SSPro and JNet and evaluates them. These three different algorithms are different methods of assigning a fold to a polypeptide sequence on the basis of already solved crystallographic structures deposited in the Protein Data Bank. Phyre server collects scores and evaluates these predictions to give the consensus assignment. On the basis of this analysis the most likely fold of *S. aeruginosa* laccases was assigned and the best possible modelling templates were selected. Results of fold recognition computation are presented in Figure 7.5 and 7.6.

From a selection of structures available; three were selected for the modelling process on the basis of their predicted fold similarity score:

- 1GYC – the crystal structure of *Trametes versicolor* laccase containing a full complement of coppers at the resolution of 1.90 Å (Piontek et al., 2002). This structure shows 61 % sequence identity and 71 % similarity with N1C2 laccase. It also exhibits 66 % sequence identity and 75 % similarity with N2C1 protein.
- 1HFU – the crystal structure of 'type 2 Cu-depleted' *Coprinus cinereus* laccase at the resolution of 1.68 Å (Ducros et al., 2001). This structure shows 56 % sequence identity and 70 % sequence similarity with N1C2 laccase. It also exhibits 62 % sequence identity and 73 % sequence similarity with N2C1 protein.

- 2HRG – the crystal structure of *Trametes trogi* blue laccase complexed with p-methylbenzoate at the resolution of 1.58 Å (to be published). This structure shows 61 % sequence identity and 70 % sequence similarity with N1C2 laccase. It also exhibits 66 % sequence identity and 75 % sequence similarity with N2C1 protein.

After selection the structural alignment of these laccases and the protein sequence of each of the *S. aeruginosa* proteins was performed. Information provided by the multiple sequence alignment was used to guide the placement of insertions and deletions. In other words the placement of insertions and deletions was guided by known crystallographic structure; indels were shifted towards loop regions if indel placement incoherence was observed. Another guideline for the structural alignment implies that conserved fragments of internal regions of proteins should not be realigned, even if the amino acid sequence suggests otherwise. The structural alignment files were prepared with DeepView (Guex and Peitsch, 1997) and were submitted for modelling using two different modelling engines:

- SWISS-MODEL – this engine transfers the peptide backbone template(s) coordinates into model coordinates and models all insertions and deletions according to internal databases; energy minimization algorithms and rotamer databases are used for side chain modelling (Guex and Peitsch, 1997)
- MODELLER – this engine uses the template structures differently, the main principle of modelling is satisfaction of spatial restrains of the created model, the algorithm uses template(s) coordinates randomizes them up to the value of 4Å and tries to select the output that satisfies those constraints as well as possible (Sali and Blundell, 1993)

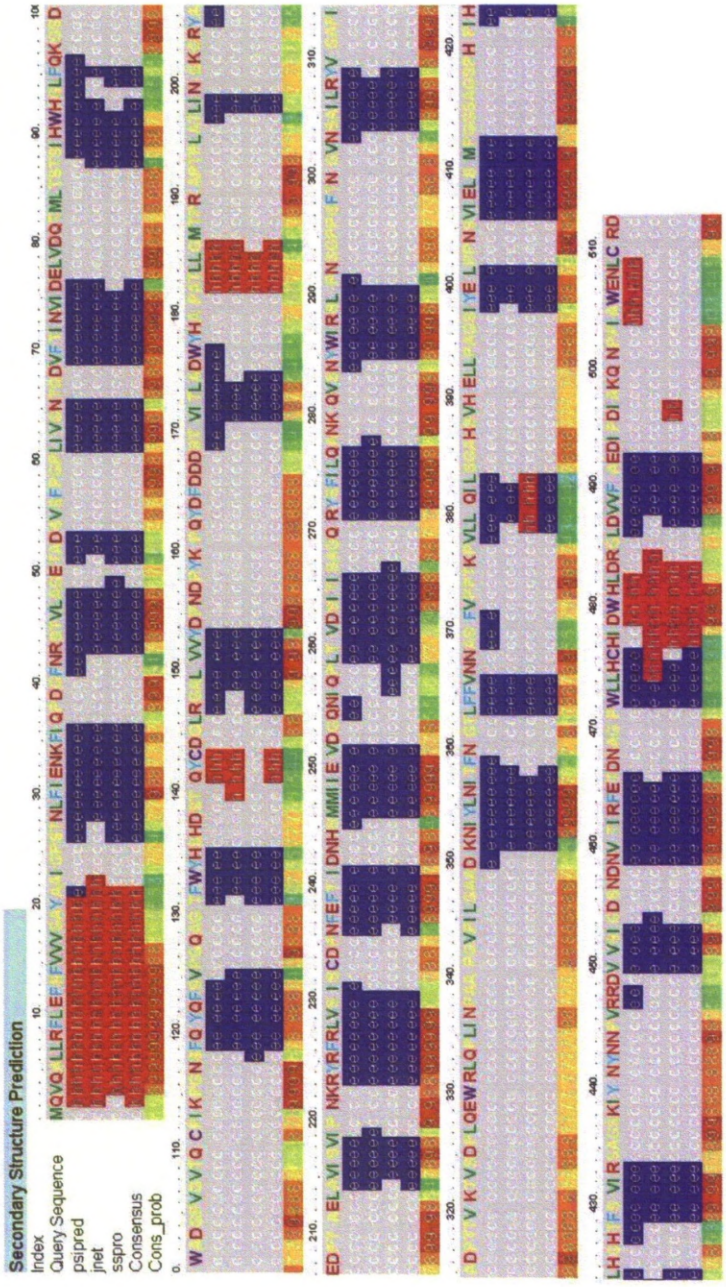


Figure 7.5 Fold recognition results of N1C2 laccase obtained with Phyre.

Helical regions are indicated in red and a letter "h"; extended strand regions are indicated in blue and a letter "e"; random coil is indicated in grey and a letter "c". Psipred, jnet and sspro are computational methods; consensus is the most probable outcome of the prediction; cons_prob is the probability of prediction (0-lowest, 9 highest)

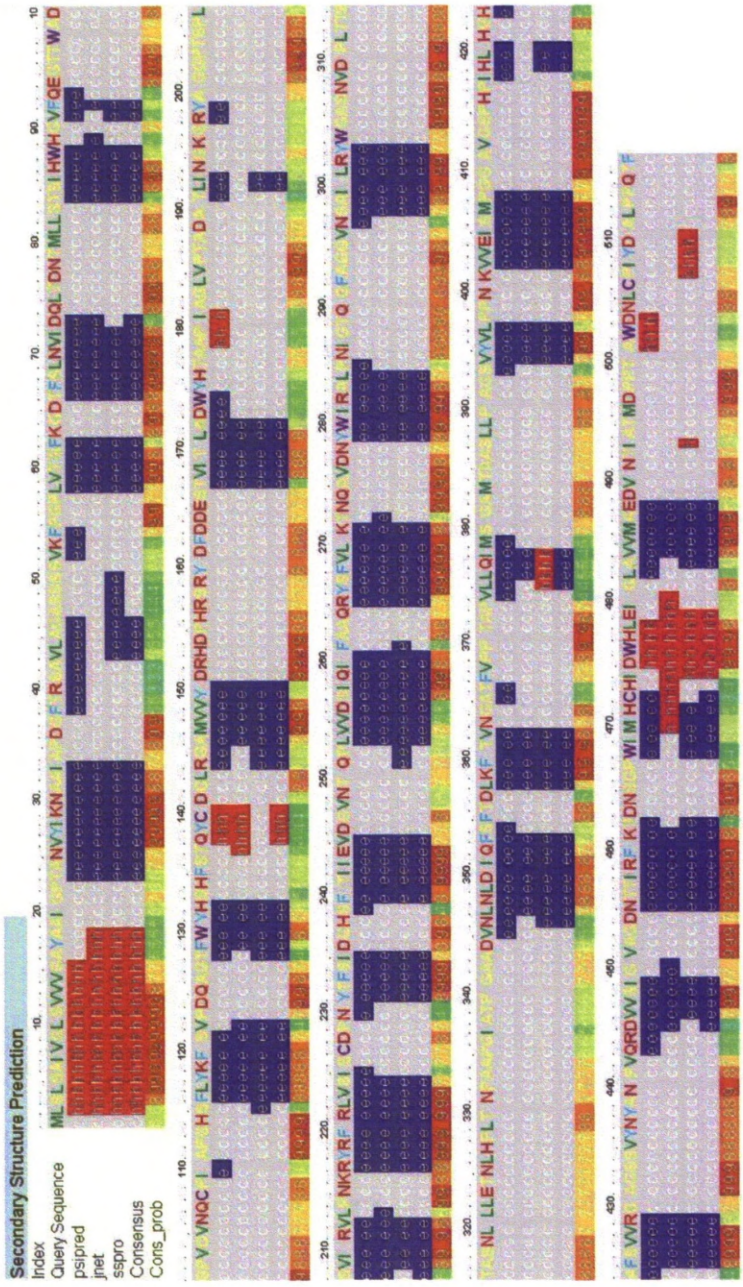


Figure 7.6 Fold recognition results of N1C2 laccase obtained with Phyre.

Helical regions are indicated in red and a letter "h"; extended strand regions are indicated in blue and a letter "e"; random coil is indicated in grey and a letter "c". Psipred, jnet and sspro are computational methods; consensus is the most probable outcome of the prediction; cons_prob is the probability of prediction (0-lowest, 9 highest).

For model creations these two methods were used simultaneously, once first generation of models was obtained they were scored with Verify3D that assesses the probability that a given residue can occupy its 3D environment at given solvent accessibility (Eisenberg et al., 1997, Lüthy et al., 1992). These scores were compared and evaluated. The final model construction comprised of a number of cycles of model creations, evaluations, manual realignments of poorly scored regions and merging of best-scored fragments (Kosinski et al., 2003). The process was stopped when no manipulation could improve the total model score without seriously compromising the local score.

5.2. Overall architecture of *S. aeruginosa* laccases.

Analysis of the molecular models of *S. aeruginosa* laccases supplemented with the data provided from MSA and fold recognition computation suggest that both enzymes of *S. aeruginosa* have a structure very similar to other fungal laccases. The models have a typical size of around 500 amino acids arranged in 3 β-barrel domains. The T1 copper site is situated in domain C that is built from amino acids 303 – 479 of both enzymes. Copper sites T2/T3 are situated at the interface of all three domains with coordinating amino acids being a part of domains A and C (all numbering refers to mature proteins without signal peptide). The overall architecture is shown in Figure 7.6 and contains four copper atoms with all amino acids coordinating them, as well as any sugar molecules bound to laccase N1C2. The figure below summarises those findings. Domain A is marked in green, domain B in brown and domain C in pink. Inter domain copper cluster T2/T3 is visible in the centre of the scheme and shown in grey whereas catalytic copper T1 is presented in dark blue. Both models seem to have two disulfide bridges conserved among all fungal laccases:

- The disulfide bridge C112-C212 for laccase N1C2; C122-C209 for laccase N2C1 stabilizes the domain A with respect to domain B
- The disulfide bridge C90-C489 for laccase N1C2; C90-C488 for laccase N2C1 connect two parts of domain A, one that begins with N-terminus of the protein with the last α helix at the C-terminus of the polypeptide chain

These two disulfide bridges are responsible for the proper orientation of T2/T3 copper plane bound to domain A with respect to catalytic T1 copper of domain C (Piontek et al., 2002). Unfortunately there are no mutagenesis studies in related proteins that could confirm this hypothesis. All disulphide bridges are highlighted in Figure 7.7.

Copper T1 is coordinated by two histidine residues and a cysteine that are a part of the well conserved motifs described above. The coordination of the trinuclear copper cluster also reflects the information obtained from multiple sequence alignment and was summarized in Table 7.1. In the case of the glycosylated laccase N1C2 four putative glycosylation sites were filled with N-acetylglucosamine molecules as in homologous 1GYC structure that was solved with full carbohydrate content. The molecular model shows that sugar distribution of N1C2 might be very uneven with most of the carbohydrate molecules being bound to a very small patch of domain C.

The highly conserved regions of laccases that could not be analyzed in sufficient detail at the raw amino acid sequence level were reviewed after constructing the molecular models of both laccases. The analysis shows that those regions are usually central parts of β sheets of individual domains, and do not seem to have any apparent effect on catalytic residues but rather maintain the overall fold of the protein. Other well-conserved patches of amino acids contain well conserved disulphide (described above) or salt bridges connecting individual domains with one another.

The models suggest that molecular surface of both proteins might differ significantly with respect to the charge distribution. The surface of laccase N1C2 is very negatively charged, whereas N2C1 seems to be more neutral. When we consider the ratio of negatively charged residues (Asp + Glu) to positively charged residues (Arg + Lys) in both proteins it is 48:29 and 38:26 respectively, the most of negatively charged are surface exposed as expected from native isoelectrofocusing experiments from Chapter 4; Figure 7.8 presents the molecular surface of both proteins with an indication of the catalytically important sites.

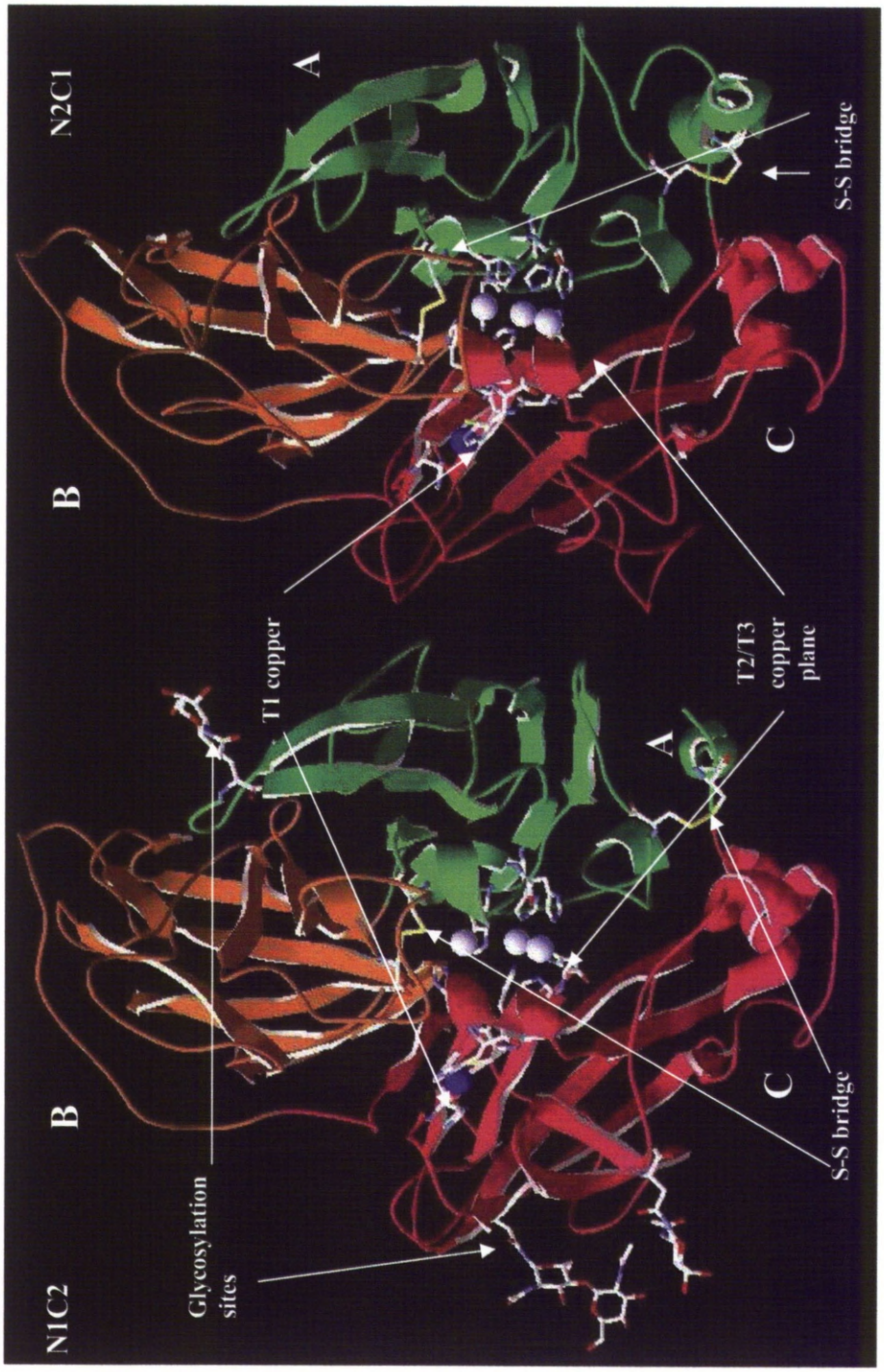


Figure 7.7 The overall architecture of *S. aeruginosa* laccases.

Domains are highlighted in green (domain A), brown (domain B) and pink (domain C). Arrows indicate S-S bridges, catalytic centre and glycosylation site. Figure made with Deep View.

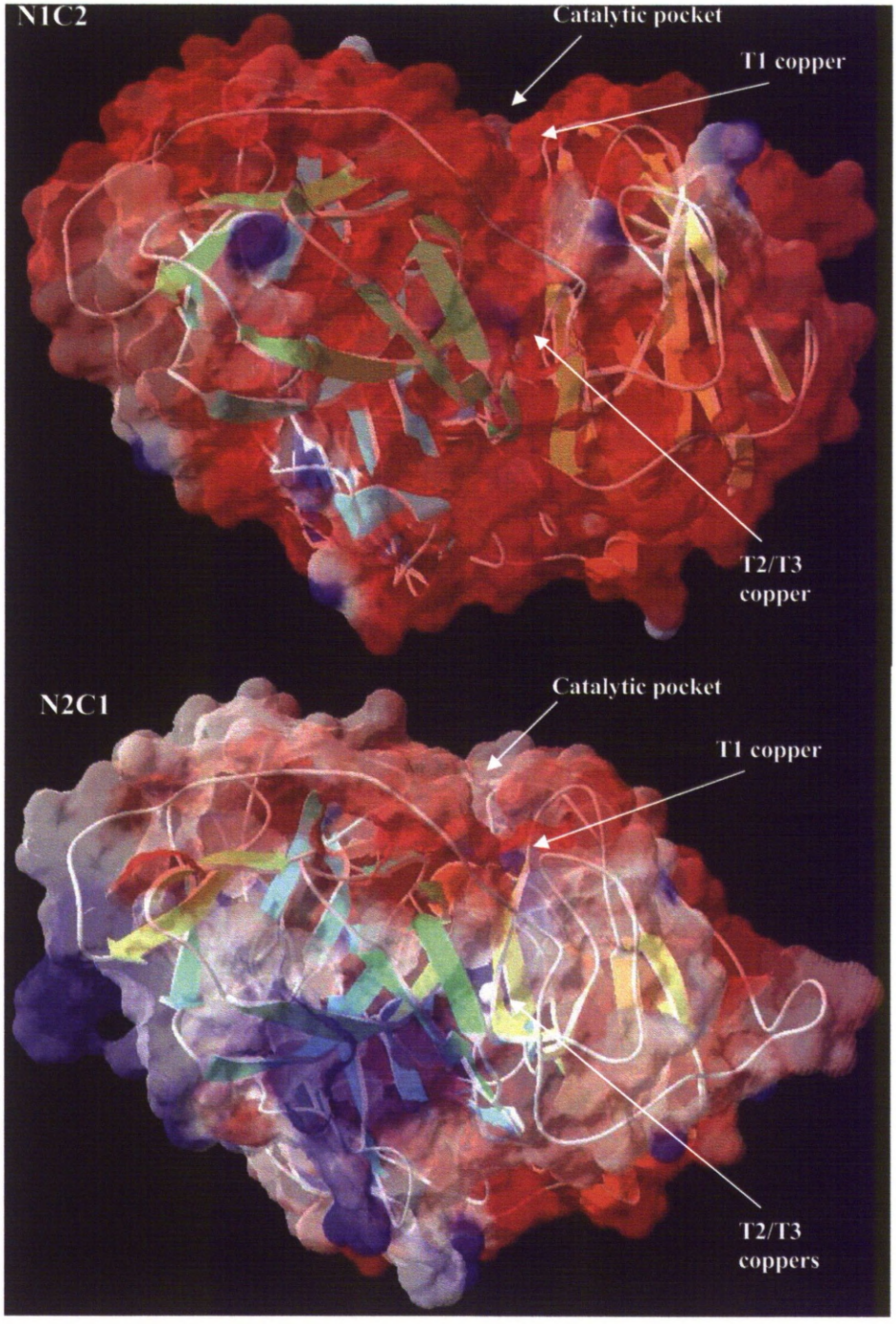


Figure 7.8 Molecular surfaces of N1C2 and N2C1 laccases.

Negative charge (< -1.800 kT/e) is shown in red, positive charge (> 1.800 kT/e) is presented in blue. Arrows indicate catalytically important sites. Figure made with Deep View

5.3. Catalytic site of *S. aeruginosa* laccases.

The catalytic sites of *S. aerugnosa* laccases do not seem to differ a lot from other laccases as the overall architecture is conserved. The entrance to the catalytic pocket contains an α-helix that plays a pivotal role in enzyme function by providing a copper coordinating histidine, an axial ligand cysteine and isoleucine in place of second axial ligand on two β strands surrounding this helix. Three other structures that surround the catalytic pocket comprise a well-conserved β hairpin with long loop that coordinates T1 copper through another histidine residue, and two structures of higher sequence variability among organisms (Figures 7.10 and 7.11).

The first of these structures is the outmost β hairpin that contains two well conserved phenylalanine residues in the centre of the hairpin and a very high variability of length and type of amino acid in the loop region. Another loop located on the opposite side of the pocket is conserved in length but its composition varies significantly among laccases. It is believed that these two structures could have importance for the substrate specificity of the laccases. The schematic drawing on Figure 7.6 presents copper T1 coordinated with two histidines and a cysteine, the visualisations of catalytic site are presented on Figure 7.7 and 7.8.

The T2/T3 copper sites represent a standard planar trinuclear site that is characteristic of laccases (Piontek et al., 2002). It contains two atoms of triple histidine coordinated copper and one copper atom coordinated by only two histidines. All these residues lie within well-conserved motifs and are almost 100 % identical in all laccases; they were previously summarized in Table 7.1.

5.4. The proposed effects of T1 copper environment on redox potential and yellow character of *S. aeruginosa* laccases.

Two histidines lie on opposite sides of the T1 copper; one is a part of a long loop, the other is a part of an α-helix. The metal atom is also coordinated through a highly conserved cysteine residue. Unlike other copper oxidases that have tetra coordinated T1 copper with methionine as fourth coordinating ligand (Giardina et al., 2010) there is no coordinating amino acid on the opposite side of the metal. Laccases have a hydrophobic amino acid in this position at a longer distance that does not coordinate the copper atom (Piontek et al., 2002). It is believed that this amino acid can have significance for the redox potential of laccases and leucine was found conserved among some of the high redox potential laccases (Rodgers et al., 2009). The previously mentioned α-helix containing the metal coordinating histidine is also believed to have significant influence on the redox potential of laccases due to the postulated possibility of helix withdrawal from catalytic site through a hydrogen bond that causes T1 Cu-His bond elongation and charge displacement, favourable for high redox catalysis (Piontek et al., 2002). A schematic drawing that facilitates analysis is presented on Figure 7.9. To further help to interpret these figures, corresponding amino acid residues are summarised in Table 7.2 and marked with arrows on Figures 7.12 and 7.13. Figure 7.10 presents crystal structures of known blue laccases; Figure 7.11 presents modelling studies of *S. aeruginosa* laccases. Figures 7.12 and 7.13 present corresponding multiple sequence alignments.

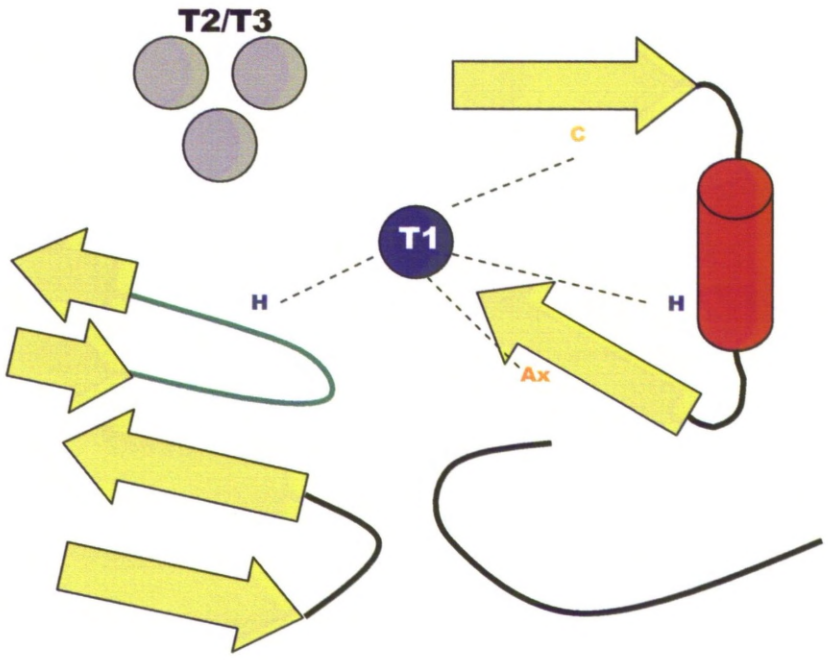


Figure 7.9 Schematic drawing of catalytic site of *S. aeruginosa* laccases and T1 copper environment.

T1 – T1 copper atom
T2/T3 – T2/T3 copper atoms
C – cysteine coordinating copper T1
H – histidine coordinating copper T1
Ax – axial copper ligand (Leu)
Green loop is the loop of variable aminoacid composition highlighted in green in Figures 7.10-6.14

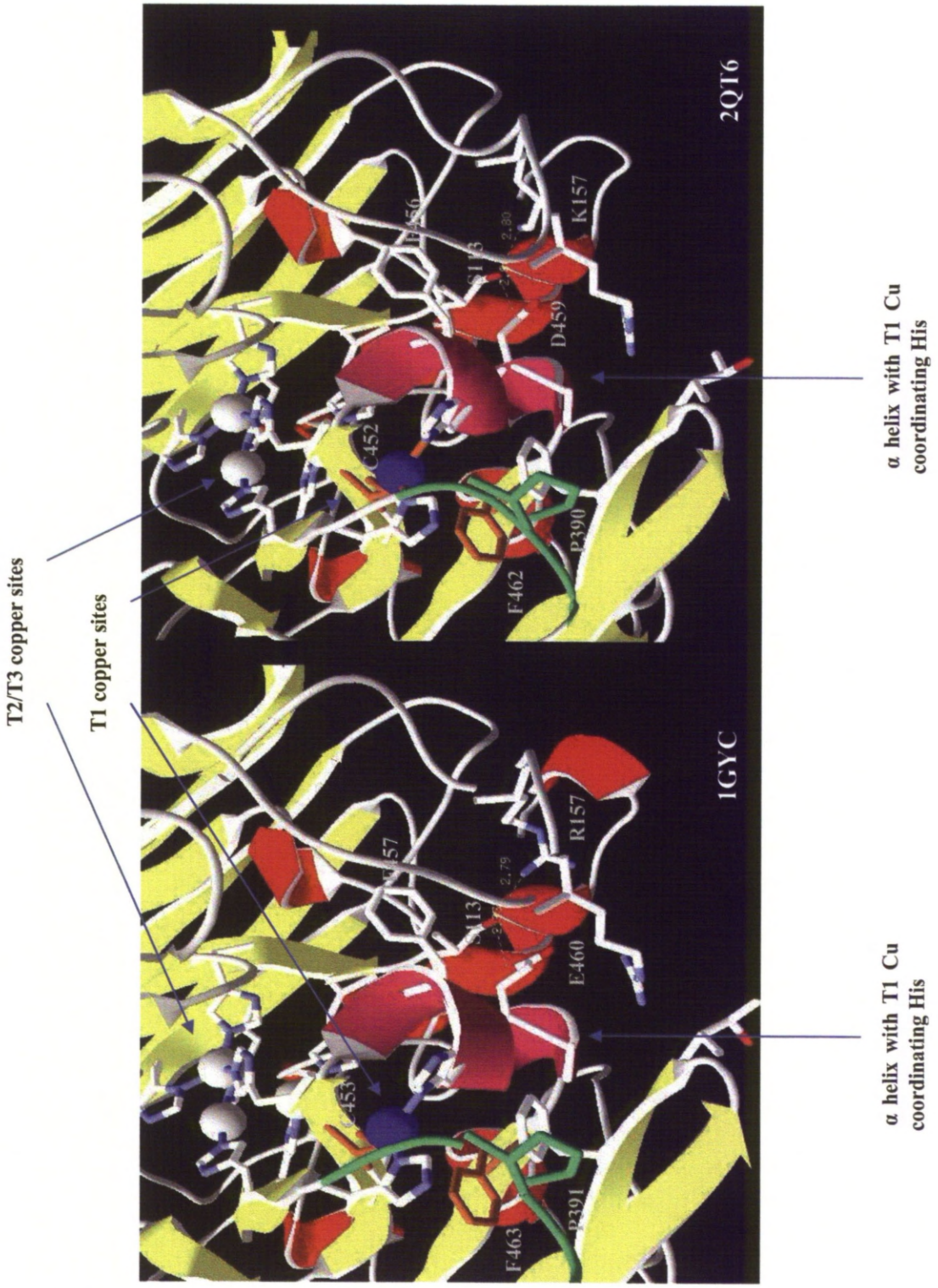


Figure 7.10 Catalytic pocket of blue laccases of *Trametes versicolor* (PDB ID: 1GYC) and *Lentinus tigrinus* (PDB ID: 2QT6).

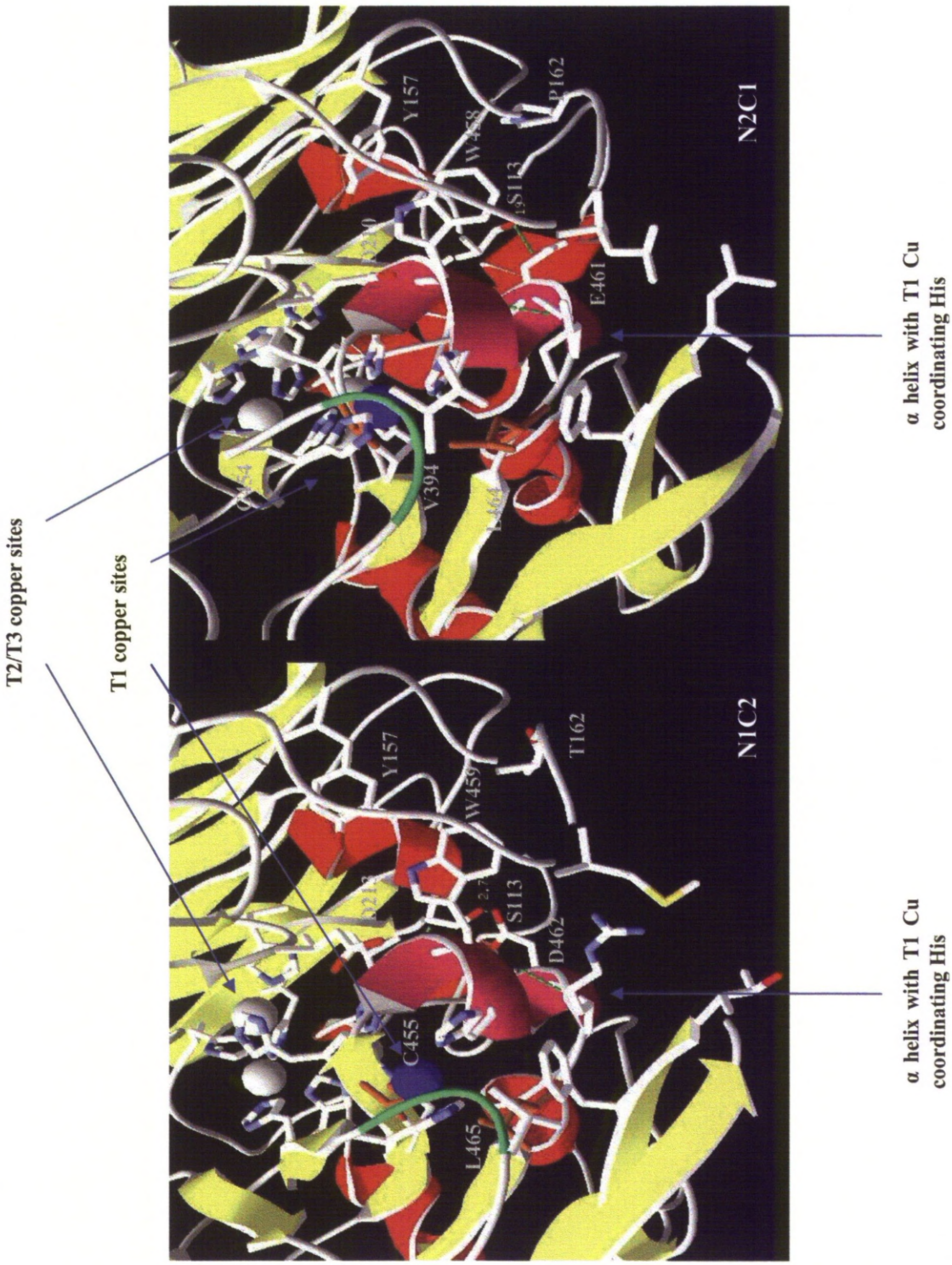


Figure 7.11 Catalytic pocket of yellow laccases of *Stropharia aeruginosa* N1C2 and N2C1.

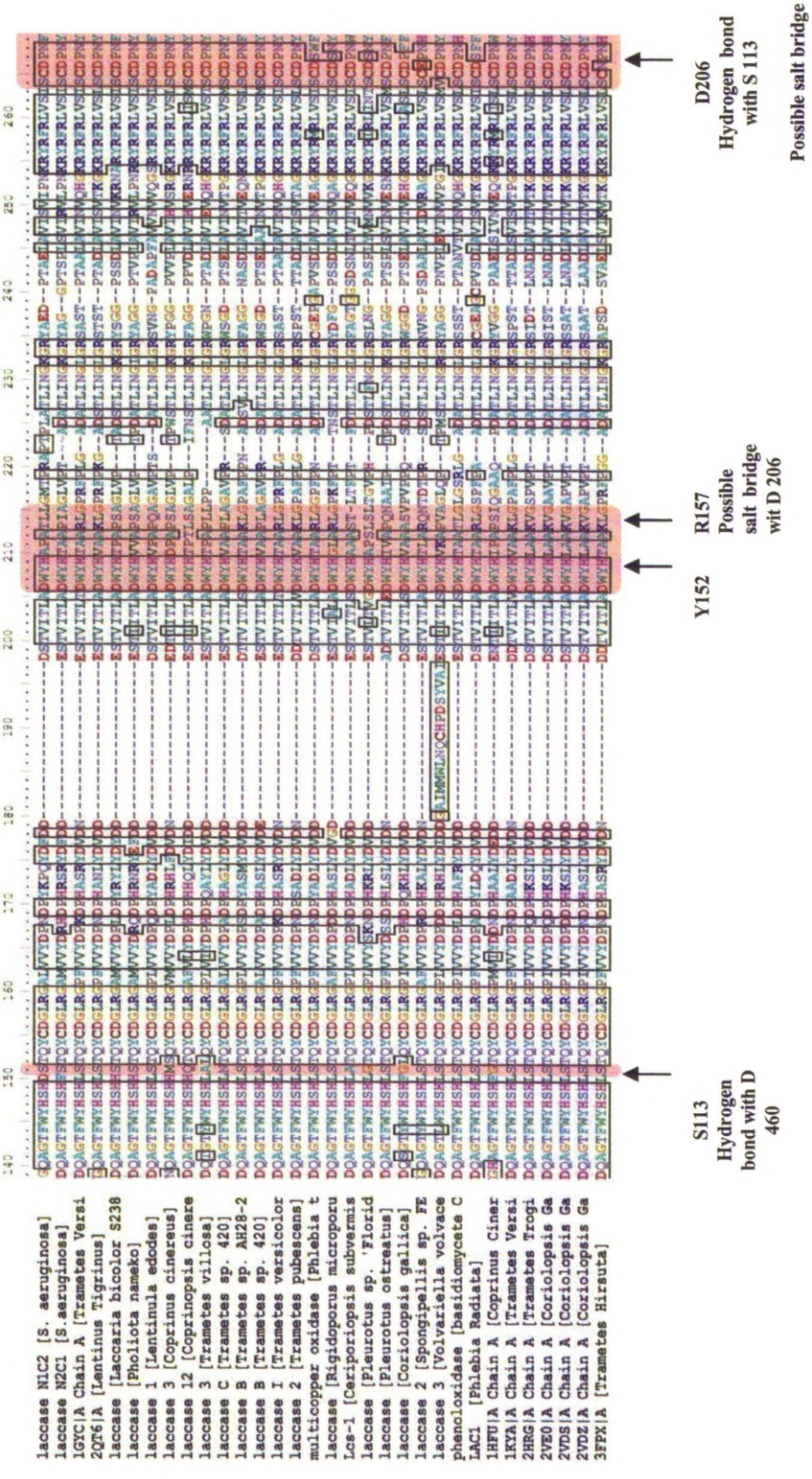


Figure 7.12 Multiple sequence alignment of selected fungal laccases showing conservation of residues presented on Figures 7.10 and 7.11.

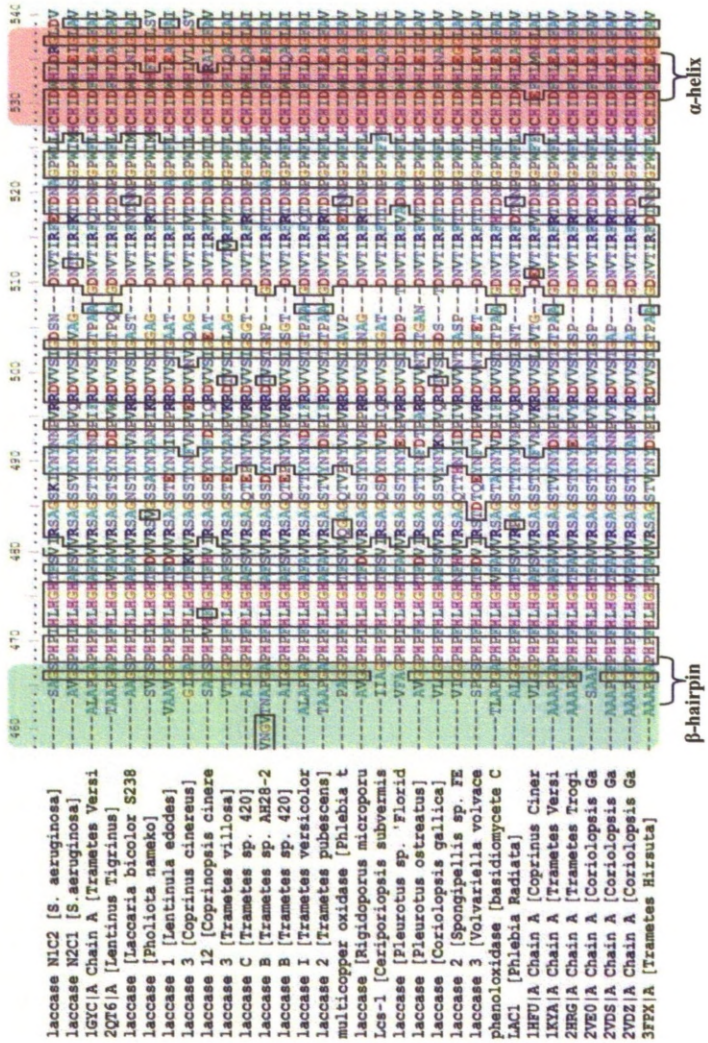


Figure 7.13 Multiple sequence alignment of selected fungal laccases showing conservation of residues presented on Figures 7.10 and 7.11.

Identical aminoacids are marked with boxes. Green shading presents variable loop region indicated in green on Figures 7.9 - 7.11. Red shading is an α-helical region.

Table 7.2 Important residues surrounding the catalytic site of selected fungal laccases and their proposed functions. 1GYC - *Trametes versicolor* laccase; 2QT6 - *Lentinus tigrinus* laccase; N1C2, N2C1 - *S. aeruginosa* laccases.

1GYC	**2QT6**	**Function**	**N1C2**	**N2C1**	**Function**
S113	S113	Hydrogen bond formation	S113	S113	Hydrogen bond formation
Y152	Y152		Y157	Y157	Hydrogen bond formation
R157	K157	Salt bridge formation	T162	P162	
D206	D206		D213	D210	Hydrogen bond formation
P391	P390	unknown	A395	V394	unknown
C453	C452	T1 ligand	C455	C454	T1 ligand
F457	F456		W458	W459	Hydrogen bond formation
E460	D459	Salt bridge formation	D462	E461	Hydrogen bond formation
F463	F462	T1 axial amino acid	L465	L464	T1 axial amino acid

To analyze the catalytic site of what appear to be the yellow laccases from *S. aeruginosa* it was decided to compare them with the two blue laccases wich have the best known structural information. The two structures selected for the comparison were the crystal structure of *Trametes versicolor* laccase deposited in PDB with the accession number 1GYC. This structure contains the full complement of copper atoms and an array of literature information (Piontek et al., 2002). The second selected

structure was the blue laccase from *Lentinus tigrinus*, this structure (PDB ID: 2QT6) also contained a full complement of copper atoms, and possessed a significant amount of literature data (Ferraroni et al., 2007). The essential information about this protein is that there are reports that the fungus is capable of producing both blue and yellow laccases when grown on different media (Leontievsky et al., 1997b). The corresponding author of these publications reports that yellow and blue laccases are two different forms of the same protein that can be interconverted, but so far no further information is known (Leontievsky, personal communication).

Here it is attempted to discuss this hypothesis through analysis of molecular models. It should be noted that homology modelling of protein structures, like any other computational method, posses its limitations, mainly in assigning the side chain conformations. The subsequent analysis of hydrogen bonds and salt bridges associated with protein movement is a hypothesis. Thus results and interpretations should be treated carefully as a hypothesis requiring detailed confirmation by crystallographic and NMR studies as it involves a dynamic system. All predictions of hydrogen bonds and salt bridges were performed with WHATIF (Vriend, 1990).

In the paper from 2002 Piontek (Piontek et al., 2002) (structure 1GYC) proposed that the influence of α-helix movement caused by a hydrogen bond created between serine 113 and glutamic acid 460 of this helix, which creates T1-Cu bond elongation and a favourable microenvironment for high redox catalysis (presented on Figure 7.14). Unfortunately, since then, no further confirmation of this hypothesis has been made. This is only one possible mechanism, that could lead to a change in orientation of this α-helix, other one occurs *via* a putative salt bridge formed between glutamic acid 460 and arginine 157 shifting the helix "sideways" from its original position (presented on Figure 7.14) The aforementioned serine 113 could also play its part in this coordination. Interestingly all the structures deposited currently in the PDB contain this acidic-basic ion pair in these positions. Laccases of *S. aeruginosa* do not

seem to have this pair present; whereas an acidic residue in the helix remains conserved the basic residue is substituted by threonine and proline. The computation showed a possible hydrogen bond between tryptophan of N1C2 (W458) and N2C1 (W459) of this α-helix that replaces phenylalanine found in all crystal structures and a conserved tyrosine (Y157) - aspartic acid (D213, D210) pair that can form different hydrogen bonding network than the one proposed by Piontek (Piontek et al., 2002). The effect of this coordination would be different positioning of this α-helix in structures of *S. aeruginosa* laccases and maybe in all yellow laccases. These considerations are visualized on Figure 7.15. At the moment the two *S. aeruginosa* protein sequences are probably the only known sequences of yellow laccases (Leontievsky, personal communications), thus more research is needed to answer this question.

When analysing the entrance to the catalytic site one may observe that the β-hairpins (indicated in green on Figure 7.10 and 7.11) are two amino acids shorter in yellow laccases than they are in the structures of their blue counterparts. In addition, the yellow laccases do not contain proline residue on the β-hairpin turn, they do however posses serine residues at position -2 to this turn in the direction of T1 copper. These changes at the β-hairpin turn might contribute to the hydrogen bonding of an unknown mediator that is believed to be responsible for the yellow colour of these proteins (Leontievsky et al., 1997b). The contribution of these two factors could provide more space for this unknown mediator to be permanently bound to the copper T1 site reducing it and quenching the characteristic blue colour of yellow laccases without affecting the catalytic activity.

When we expand these considerations to the partial sequences of the *S. aeruginosa* laccases one can observe that these findings are also confirmed by the partial sequences of *S. aeruginosa* laccases. All partial protein sequences lack the putative salt bridge (Figure 7.16, blue shading) and posses all three amino acids that can form an alternative hydrogen bonding (Figure 7.16, red shading). Also the β-hairpins (indicated in green

on Figure 7.16) are of the same length and lacking the proline residue similar to that observed in the sequences of N1C2 and N2C1. Some of these loops contain a serine residue for the putative hydrogen binding of a mediator (N1C2, N2C1 and NMC1), whereas NMC3 and NMC4 contain a glycine residue in this position instead that increases flexibility of the loop, however cannot form proposed hydrogen bonding.

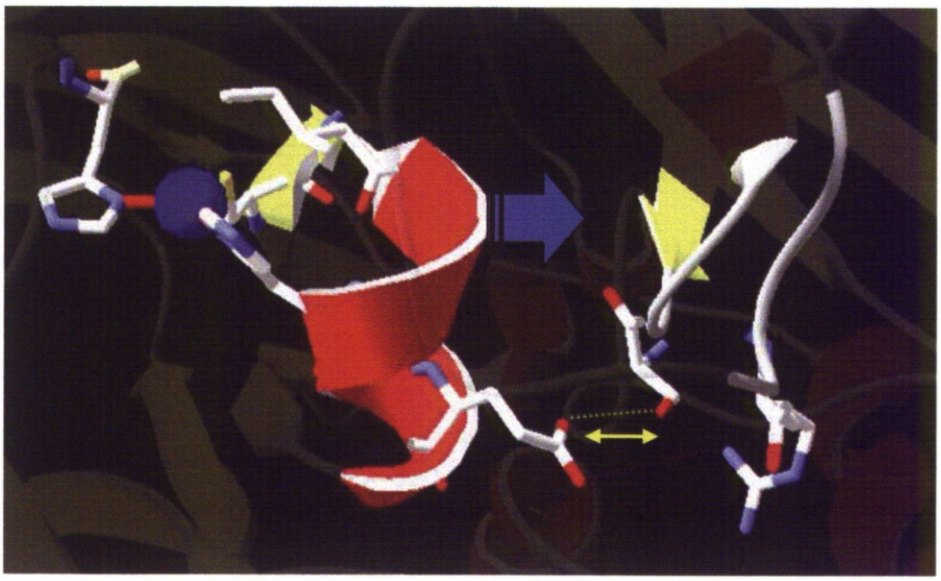


Figure 7.14 Catalytic copper T1 and neighbouring α-helix of *T. versicolor* laccase structure 1GYC.

Original hypothesis presented by Piontek et al. (2002) suggests withdrawal of an α-helix through the action of a hydrogen bond. This interaction is indicated with the yellow dotted line.
An alternative explanation is marked with magenta line – the helix withdrawal is performed through the action of a salt bridge not the hydrogen bond.
Blue arrow indicates the helix movement out of the catalytic site

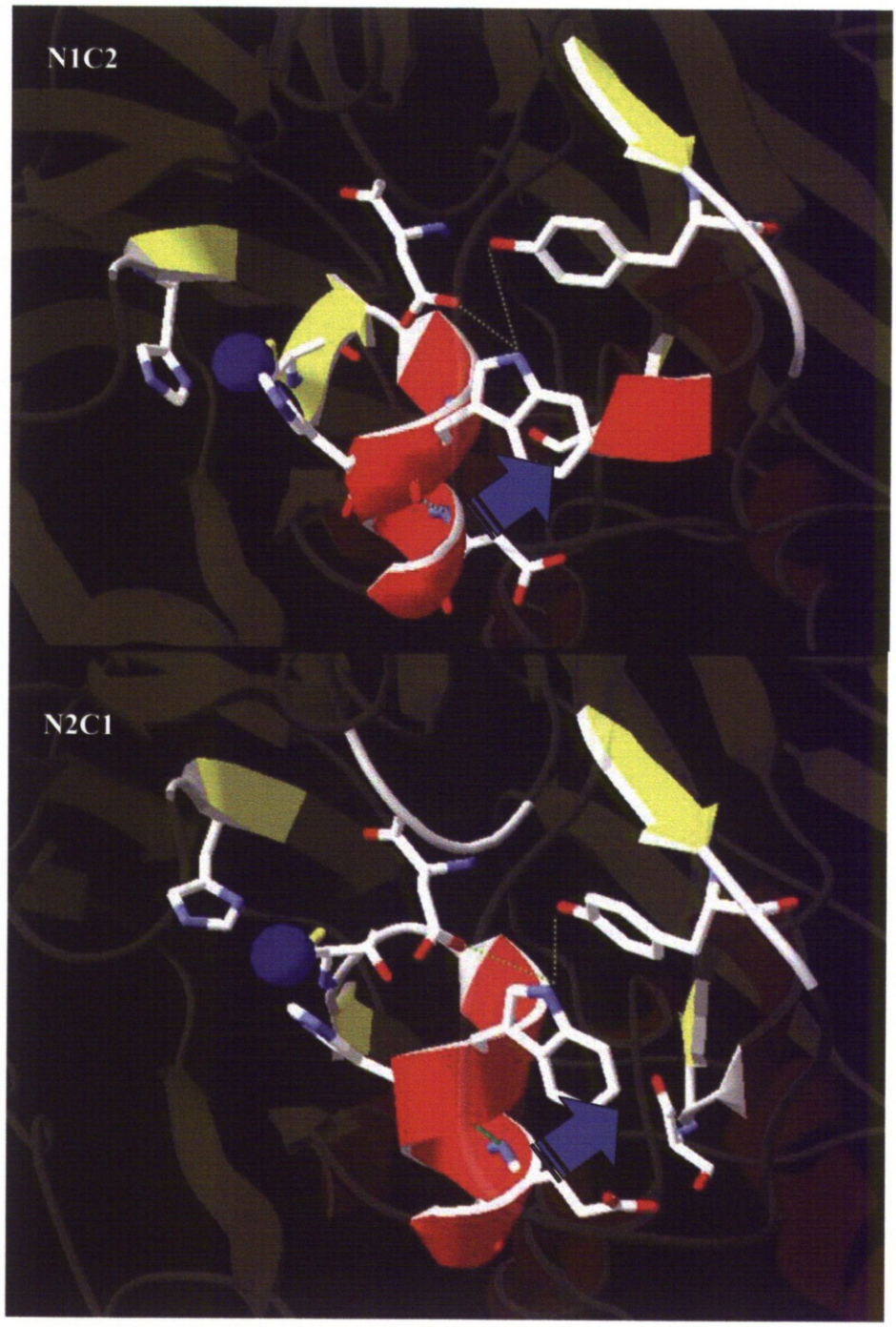


Figure 7.15 Catalytic copper T1 and neighbouring α-helix of *S. aeruginosa* models.

Potential hydrogen bonds between tryptophan, aspartic acid and tyrosine are marked with yellow dots. The blue arrow indicates helix movement through the action of a hydrogen bond

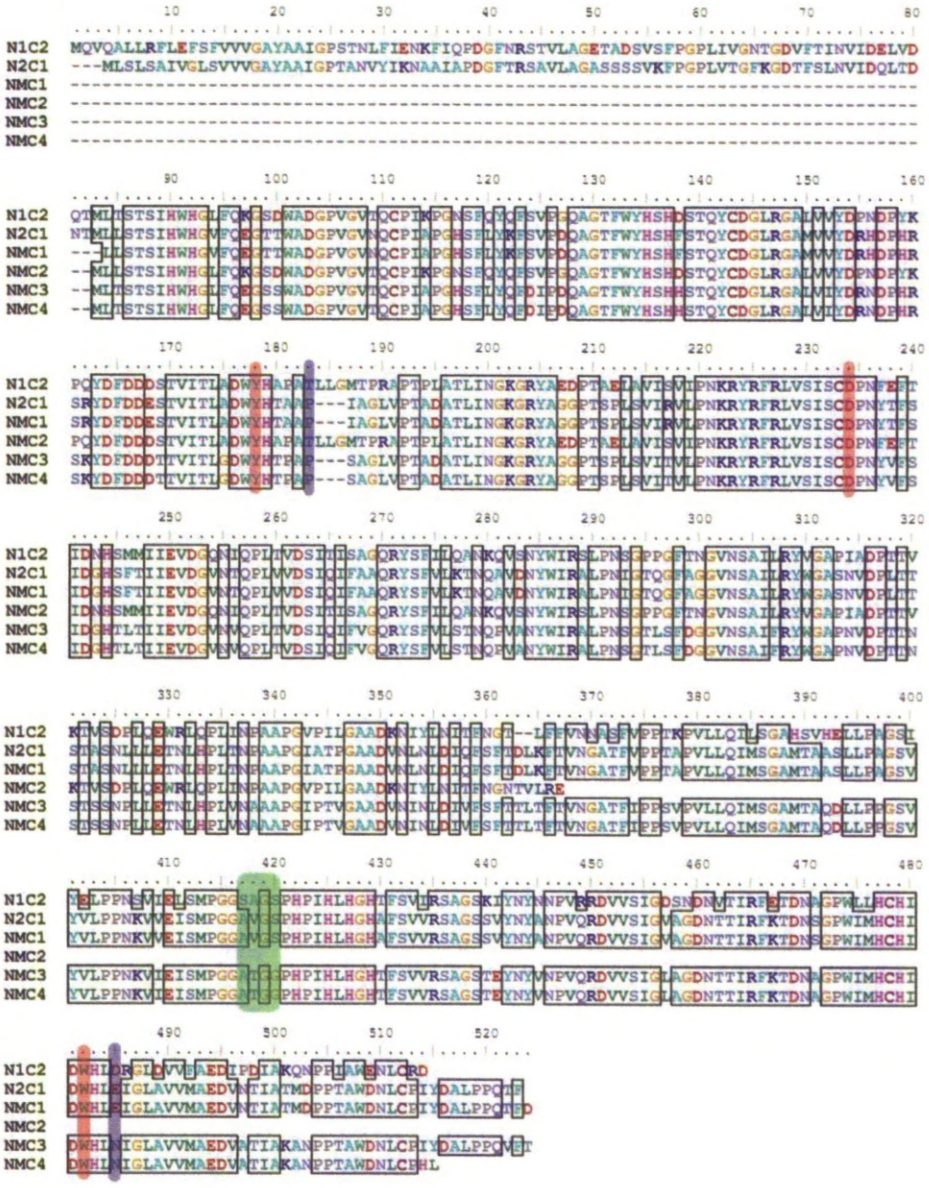


Figure 7.16 Multiple sequence alignment of *S. aeruginosa* full and partial sequences.

Full (N1C2 and N2C1) and partial (NMC1-NMC4) sequences of *S. aeruginosa* laccases were multiply aligned to verify the hypothesis of the possible alternative hydrogen bonding network surrounding T1 copper site. Amino acids not forming a putative salt bridge are shaded in blue, amino acids that contribute to alternative hydrogen binding network are shaded in red, the short loop leading to the catalytic T1 copper is shaded in green. Identical amino acids are marked in boxes.

6. Summary.

The bioinformatic analysis of *S. aeruginosa* laccases revealed both similiarities and differences between those proteins and other fungal laccases. The sequence analysis of *S. aeruginosa* proteins showed that these secreted enzymes possess 13 - 14 introns per coding sequence. One of the full length proteins contains glycosylation sequences (three out of four partial sequences also show possible glycosylation sites), consistent with results from 2D electrophoresis of flow through and eluate fractions that showed both glycosylated and non-glycosylated proteins. When sequence analysis was completed, molecular structure models were created using the FRankenstein Monster's approach (Kosinski et al., 2003). The analysis of those models revealed that although the overall architecture of those proteins is similar to other laccases, some fine details regarding the T1 copper environment may be able to explain their yellow character. The overall architecture shows that those proteins are composed of three domains, with metal atoms situated in two sites: the reduced copper cluster T2/T3 situated at the interface of three domains and catalytic T1 copper site situated in C terminal domain C. Molecular surface charge of both proteins differ significantly but is generally shifted towards the acidic range. Previous isoelectrofocusing experiments from Chapter 4 confirm these findings. A detailed analysis of the T1 copper environment of yellow laccases and two well known examples of blue laccases suggest that different bonding of an α-helix and a shorter loop leading to the catalytic site might be responsible for lack of a characteristic laccase absorbtion spectrum of these proteins.

Chapter 8

DISCUSSION

The original aim of this project was to follow Jonathan K. Moore's doctoral project "Pursuit of a novel haloperoxidase for asymmetric biocatalysis" (Moore, 2007) and identify new genes encoding haloperoxidase(s). Using Moore's guidance a basidiomycete fungus *Stropharia aeruginosa* was selected for this purpose.

To date, the best-known enzyme showing haloperoxidase activity is the chloroperoxidase of an ascomycete *Caldariomyces fumago*. This protein has been studied for over 40 years (Morris and Hager, 1966, Hager et al., 1966) and despite many limitations, including resistance to heterologous expression and vulnerability to high concentrations of peroxide, a real alternative has not been found. Basidiomycete halogenating activities were first identified in *Phanerochaete chrysosporium* (Sheng and Gold, 1997); however the best example comes from *Agrocybe aegarita* peroxidase (AaP) (Ullrich et al., 2004). However, the halogenating abilities of this enzyme were rather moderate, with AaP showing 12 % of the brominating activity of *C. fumago* chloroperoxidase (CPO) per mg of protein and 4.7 % of its chlorinating activity. Other activities were significantly higher AaP exhibited 250 times higher ABTS activity than CPO, whereas specificity towards DMP and benzyl alcohol were 10 and 50 times higher. It became apparent that haloperoxidase is only an ancillary activity of AaP this was confirmed with further findings and reclassification of AaP. Further analysis of enzyme substrate specificity (Ullrich et al., 2008) and molecular analysis showed that this enzyme is a peroxygenase not a true haloperoxidase. It shares < 30 %

sequence homology with CPO (Pecyna et al., 2009) and its crystal structure has been now elucidated (Piontek et al., 2010). Observations of other basidiomycete enzymes showing secondary haloperoxidase activities were made for *Coprinus radians* peroxygenase (Anh et al., 2007) and manganese peroxidase of *Phanerochaete chrysosporium* (Sheng and Gold, 1997) however neither of those enzymes can be regarded as a real alternative for CPO.

The initial aim was to isolate the enzyme from *S. aeruginosa*, however significant problems with contaminating pigment (probably melanin) and other technical difficulties rendered the task of isolating the *S. aeruginosa* putative haloperoxidase impossible and even raised the question as to whether such an enzyme ever existed in this organism. None of Moore's experiments showing chlorinating activities of *S. aeruginosa* extracellular extracts could be repeated using two different haloperoxidase assays: the monochlorodimedone assay (Hager et al., 1966) and thionin assay (Manoj and Hager, 2006). To address this problem a concomitant to haloperoxidase, peroxidase activity was used to follow the protein during the purification process. Once proteins with peroxidase activity were liberated from the contaminating pigment through Q Sepharose batch adsorption, they were tested for haloperoxidase activity, without success.

It was then decided to focus on enzymes showing the highest oxidase activity in the extracts of *S. aeruginosa*, laccases. As a result two laccases from an array of these enzymes secreted by *S. aeruginosa* were purified using chromatographic methods and analysed biochemically. The two proteins were denoted as N and E due to their behaviour on Phenyl Toyopearl hydrophobic interaction chromatographic medium (N – non-bound, E – eluted). Both proteins are secreted monomeric glycoproteins and showed similar molecular weight of around 55 kDa, confirmed by size exclusion chromatography and SDS-PAGE. Proteins were isolated using

ConcanavalinA affinity chromatography which suggests that they are both N-glycosylated. All these findings are very consistent with current knowledge about laccases that have been already reviewed several times (Baldrian, 2006, Giardina et al., 2010, Claus, 2004, Morozova et al., 2007). Both proteins exhibit the typical substrate specificities of laccases. Highest activity was exhibited towards the synthetic substrate ABTS and lignin building blocks, ferulic and sinapic acids. ABTS is the most widely used substrate of laccases and the typical substrate used for their screening (Alcalde et al., 2005, Chairattanamanokorn et al., 2006). Ferulic acid and sinapic acid are not as commonly used as ABTS but have been tested for several laccases and found more specific than ABTS (Chefetz et al., 1998, Garzillo et al., 1998). The laccases of *S. aeruginosa* show rather unusual substrate specificity characteristics. They are specific only for three characteristic compounds, and relatively unspecific for other commonly used substrates. This is in contrast to other laccases as they are known to be rather unspecific catalysts that accept a wide range of compounds (Baldrian, 2006, Giardina et al., 2010).

Further biochemical characterisation revealed the similarity of *S. aeruginosa* laccases to other fungal laccases of mesophilic origin. Proteins of *S. aeruginosa* exhibit their maximal activity between 40 and 50 °C and are thermostable up to 40 °C which are common values for mesophilic laccases (Rodriguez et al., 2008, Jordaan and Leukes, 2003). The pH range of activity and stability also follows trends characteristic for other fungal laccases. Optimal pH of activity depends on substrate, synthetic substrates like ABTS, TMB and *o*-dianisidine are best oxidised at low pH. The activity drops significantly above pH of 4.5 which is consistent with other enzymes (Baldrian, 2006). Lignin derived phenolics, including syringaldazine, posses a broader pH range of activity and are oxidised in pH more similar to that of the natural environment which is also typical of other proteins. (Frasconi et al., 2010, Pozdnyakova et al.,

2004, Pozdnyakova et al., 2006). The pH stability of both proteins are also similar to their counterparts in other organisms (Rodriguez et al., 2008) as these proteins are generally stable at a pH close to physiological for *S. aeruginosa* laccases it is 5.5 to 9.0. The inhibitory effects of small molecules also confirm that the proteins of *S. aeruginosa* are laccases, these enzymes are significantly inhibited by: sodium azide, halides, hydroxylamine, ascorbic acid and thiols; the same applies for many other laccases (Min et al., 2001, Garzillo et al., 1998). The laccases of *S. aeruginosa* were also tested for their ability to degrade synthetic azo- and athraquinone dyes; highly toxic dyes that are widely used in the textile industry. It was found that one of these proteins (E enzyme) could be applied in the degradation of an anthraquinone dye Remazol Brilliant Blue R. To a lesser extent both laccases showed some promise in decolourisation of other dyes, however more spectacular results can be found in the literature (Michniewicz et al., 2008, Kunamneni et al., 2008b).

After biochemical characterisation was completed, attempts were made to gather information about the protein sequences of *S. aeruginosa* proteins and the relationship between the different laccases. Separation was achieved by semi native 2D PAGE and the resultant spots were digested with specific proteases and analysed by tandem mass spectrometry. It was found that different spots result in identical/very similar peptide sequences and that, these peptides originate from proteins that show different degrees of glycosylation. Unfortunately BLAST searching could not identify them as sequences similar to known laccases. Multiple isoforms of ligninolytic enzymes are generally are well described in the literature and are generally differentially regulated in response to culture age and nutrient limitation (Lobos et al., 1994, Dittmer et al., 1997, Moldes, 2004). We believe that multiple isoforms of lignin degrading enzymes exist to address the heterogeneity of lignin.

When the relationship between the laccases was confirmed at the protein level it was decided to isolate the genes encoding the proteins of *S. aeruginosa*. A number of methodologies including classical methods of molecular biology and biochemistry were unsuccessfully attempted. These techniques included screening phage, bacterial and yeast cDNA libraries of *S. aeruginosa* with chromogenic substrates and molecular methods including degenerate-primer PCR and degenerate probe hybridisation. The lack of success using well established traditional methods prompted the search for an alternative approach. It was decided to use 454 sequencing technology to sequence the transcriptome of *S. aeruginosa* at very low coverage to identify fragments of homology to laccase sequences deposited in the NCBI database. After annotation, interesting sequences were amplified by inverse PCR performed on a plasmid library to obtain the full length sequence of the gene. Although a next generation sequencing approach has been already used for annotating transcriptomes of other organisms (Vera et al., 2008) and analysing lignocellulose degraders (Warnecke et al., 2007), we believe that our approach can become a useful tool for difficult to isolate genes. The method is relatively fast, inexpensive and does not require any prior sequence information. Genes of completely novel proteins could be also identified using this approach as the modification of the bioinformatic pipeline of this process can be expanded from BLAST to fold recognition techniques and other algorithms. Using this approach two full length, and many partial sequences were identified, none of which were obtained using classical methods. A putative sequence of a haloperoxidase-like protein was also identified; however similar doubts to *A. aegarita* protein regarding the substrate specificity can be raised as BLAST analysis showed sequence similarity to both CPO and aromatic peroxygenases. Enzyme isolation and/or heterologous expression are the only way to analyse if any of these activities is exhibited by the protein. So far PCR amplification of the full length gene has not been

successful, colony screening is recommended if any further plans to analyse this protein are envisaged.

During biochemical characterisation of the proteins, both enzymes of *S. aeruginosa* exhibited unusual spectroscopic properties suggesting that those proteins belong to the family of yellow laccases due to the lack of characteristic absorbance peak at 600 nm. Classical laccases are enzymes with a characteristic blue colour caused by the coordination between cysteine and an oxidized copper at site T1 (Bertrand et al., 2002). Up to now several enzymes showing other characteristics of laccases but lacking this property have been identified, and several attempts to explain this phenomenon have been proposed. The most obvious explanation is that such enzymes do not exist and yellow forms of laccases are laccases stripped of T1 copper (Martins, L.O. personal communication). If this was the case, yellow laccases should not have comparable activities to blue laccases. It is likely they would not have any catalytic activity at all as T1 copper is the catalytic copper of the multicopper oxidases. The contrary was proven by this work and other publications where yellow laccases perform equally well, or even better than blue laccases as they are capable of oxidizing non-phenolic lignins that are outside the substrate specificity of blue enzymes (Pozdnyakova et al., 2004). Another possible explanation of this phenomenon is that the products of lignin decomposition from lignin-rich growth medium can obscure the real colour of blue copper. This argument has been disputed (Leontievsky et al., 1997b) by analysing spectra of yellow laccases purified to electrophoretic homogenity. It confirmed that the T1 copper signal of yellow laccases is absent not obscured. Evidence from *S. aeruginosa* proteins suggest that lignin-derived products cannot be responsible for obscuring the blue colour, as the fungus was grown on synthetic medium without any addition of lignin. Another piece of evidence from Leontievsky's work that supports the hypothesis that yellow laccases do exist and has confirmed that both blue and yellow

laccases posses four copper atoms per molecule of enzyme (Leontievsky et al., 1996). An alternative explanation regarding the yellow colour of these proteins was drawn, it has been postulated that a mediator molecule, being an aromatic product of lignin degradation, is bound to the active site of the enzyme reducing its T1 copper and causing the disappearance of the blue colour (Leontievsky et al., 1997b). Such molecule would then act as an electron transfer mediator, just like mediators in the laccase-mediator system, but permanently bound to the enzyme. The mediator could then be oxidised more than once, achieving higher redox potential and expanding the substrate specificity to non-phenolic lignins (Kunamneni et al., 2008a). Some results from *S. aeruginosa* support this hypothesis; however the origin of this reported mediator does not have to be lignin. We believe that the mediator could be an aromatic molecule synthesised by the fungus during its growth.

To analyse the molecular basis of the mediator phenomenon, comparative models of *S. aeruginosa* proteins were constructed on the basis of structures of known blue laccases to unveil the possible differences between these proteins. Molecular model analysis suggested that *S. aeruginosa* might have an alternative pattern of hydrogen and salt bridge bonding around the central α-helix donating T1 copper-coordinating histidine. Previous analysis of 1GYC structure of *T. versicolor* laccase (Piontek et al., 2002) suggests that there exists a hydrogen bond that can withdraw this α-helix from the T1 site to create a favourable microenvironment for high redox catalysis. We believe that a similar effect can be obtained with the action of a neighbouring salt bridge, although such coordination does not seem to exist when analysing comparative models of *S. aeruginosa* laccases. Instead the α-helix seems to contain a tryptophan which is absent in the structures of known blue laccases. It is possible that this tryptophan forms a different hydrogen bond with a highly conserved tyrosine and aspartic acid. Such a hydrogen bonding network

will move the α-helix sideways from the T1 site leaving more space for a putative mediator. To date there are no other sequences of yellow laccases that could help to verify these hypothesis. In addition to this, diffracting crystals of yellow laccases have not been obtained, it is postulated that the technical difficulties result from an uneven binding/heterogenicity of a mediator molecule that hinders crystal growth (Leontievsky, personal communication) so the analysis of molecular models is to date the best attempt to explain this phenomenon.

Summarising, the original aim of this work was only partially achieved, as an incomplete gene sequence of a haloperoxidase-like protein of *S. aeruginosa* was identified, and the method to isolate the full length sequence has been suggested. It is necessary to express this protein in an heterologous host (*A. niger*, *S. cerevisiae*, *P. pastoris*) to verify its substrate specificity and identify if it becomes the first real haloperoxidase of a basidiomycete fungi or peroxygenase with ancillary haloperoxidase activity, or something completely different. Important findings were made in the methodology for quickly identifying genes of biotechnologically relevant proteins having been made. This methodology has been used with laccases of *S. aeruginosa* and a *Geomyces sp. P7* lipase. The laccases of *S. aeruginosa* that became the main topic of this work for most of the time are interesting enzymes belonging to the family of laccases. They exhibit many features characteristic for the family: molecular weight, pI existence of multiple isoforms. Their substrate specificity however is more limited than it was envisaged. Unusual spectroscopic properties classify them as yellow laccases as they lack an absorption band at 600 nm characteristic of blue laccases. To explain the possible origins of this phenomenon, a bioinformatic approach was taken. Homology models of two yellow laccases (being the first protein sequences of yellow laccases) were constructed and it was suggested that yellow laccases could possess more space around the T1 copper site than blue laccases, which would facilitate

the binding of a putative mediator molecule. To verify this hypothesis site directed mutagenesis experiments are recommended, as to date crystals of yellow laccase do not diffract well enough to solve their crystallographic structure.

References

LITERATURE USED IN A COURSE OF THE PROJECT

AKTAS, N. & TANYOLAS, A. (2003) Reaction conditions for laccase catalyzed polymerization of catechol. *Bioresource Technology,* 87, 209-214.

ALCALDE, M., BULTER, T. & ARNOLD, F. H. (2002) Colorimetric assays for biodegradation of polycyclic aromatic hydrocarbons by fungal laccases. *Journal of Biomolecular Screening,* 7, 547-553.

ALCALDE, M., BULTER, T., ZUMARRAGA, M., GARCIA-ARELLANO, H., MENCIA, M., PLOU, F. J. & BALLESTEROS, A. (2005) Screening mutant libraries of fungal laccases in the presence of organic solvents. *Journal of Biomolecular Screening,* 10, 624-631.

ANH, D. H., ULLRICH, R., BENNDORF, D., SVATOS, A., MUCK, A. & HOFRICHTER, M. (2007) The coprophilous mushroom Coprinus radians secretes a haloperoxidase that catalyzes aromatic peroxygenation. *Applied and Environmental Microbiology,* 73, 5477-5485.

ANTIPOV, A. N., SOROKIN, D. Y., L'VOV, N. P. & KUENEN, J. G. (2003) New enzyme belonging to the family of molybdenum-free nitrate reductases. *Biochemical Journal,* 369, 185-189.

ARNOLD, F. H. (2001) Combinatorial and computational challenges for biocatalyst design. *Nature,* 409, 253-257.

AYALA, M., PICKARD, M. A. & VAZQUEZ-DUHALT, R. (2008) Fungal enzymes for environmental purposes, a molecular biology challenge. *Journal of Molecular Microbiology and Biotechnology,* 15, 172-180.

BALDRIAN, P. (2006) Fungal laccases - occurrence and properties. *Fems Microbiology Reviews,* 30, 215-242.

BALLAND, A., COURTNEY, M., JALLAT, S., TESSIER, L.-H., SONDERMEYER, P., DE LA SALLE, H., HARVEY, R., DEGRYSE, E. & TOLSTOSHEV, P. (1985) Use of synthetic oligonucleotides in gene isolation and manipulation. *Biochimie,* 67, 725-736.

BATTISTUZZI, G., BELLEI, M., BORTOLOTTI, C. A. & SOLA, M. (2010) Redox properties of heme peroxidases. *Archives of Biochemistry and Biophysics,* 500, 21-36.

BAUCHER, M., MONTIES, B., VAN MONTAGU, M. & BOERJAN, W. (1998) Biosynthesis and Genetic Engineering of Lignin. *Critical Reviews in Plant Sciences,* 17, 125-197.

BAYER, E. A. (2000) The Cellulosome. IN DOYLE (Ed.) *Glycomicrobiology.* Kluwer Academic.

BAYER, E. A., BELAICH, J. P., SHOHAM, Y. & LAMED, R. (2004) The cellulosomes: Multienzyme machines for degradation of plant cell wall polysaccharides. *Annual Review of Microbiology,* 58, 521-554.

BEGUIN, P. & LEMAIRE, M. (1996) The cellulosome: An exocellular, multiprotein complex specialized in cellulose degradation. *Critical Reviews in Biochemistry and Molecular Biology,* 31, 201-236.

BENDTSEN, J. D., NIELSEN, H., VON HEIJNE, G. & BRUNAK, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol,* 340.

BERTRAND, T., JOLIVALT, C., BRIOZZO, P., CAMINADE, E., JOLY, N., MADZAK, C. & MOUGIN, C. (2002) Crystal structure of a four-copper laccase complexed with an arylamine: Insights into substrate recognition and correlation with kinetics. *Biochemistry,* 41, 7325-7333.

BLEVE, G., LEZZI, C., MITA, G., RAMPINO, P., PERROTTA, C., VILLANOVA, L. & GRIECO, F. (2008) Molecular cloning and heterologous expression of a laccase gene from Pleurotus eryngii in free and immobilized Saccharomyces cerevisiae cells. *Applied Microbiology and Biotechnology,* 79, 731-741.

BOMMARIUS, A. S. & RIEBEL, B. (2004) *Biocatalysis: Fundamentals and Applications* Wiley VCH

BOONYAKAMOL, A., IMAI, T., CHAIRATTANAMANOKORN, P., HIGUCHI, T. & SEKINE, M. (2009) Key Factors Regarding Decolorization of Synthetic Anthraquinone and Azo Dyes. *Applied Biochemistry and Biotechnology,* 158, 180-191.

BRADFORD, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Analytical Biochemistry,* 72, 248-254.

BROWN, D. W., YU, J. H., KELKAR, H. S., FERNANDES, M., NESBITT, T. C., KELLER, N. P., ADAMS, T. H. & LEONARD, T. J. (1996) Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in Aspergillus nidulans. *Proceedings of the National Academy of Sciences of the United States of America,* 93, 1418-1422.

BULTER, T., ALCALDE, M., SIEBER, V., MEINHOLD, P., SCHLACHTBAUER, C. & ARNOLD, F. H. (2003) Functional expression of a fungal laccase in Saccharomyces cerevisiae by directed evolution. *Applied and Environmental Microbiology,* 69, 987-995.

BURNETTE, W. N. (1981) "Western Blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Analytical Biochemistry,* 112, 195-203.

BUTLER, M. J. & DAY, A. W. (1998) Fungal melanins: A review. *Canadian Journal of Microbiology,* 44, 1115-1136.

CHAGAS, E. P. & DURRANT, L. R. (2001) Decolorization of azo dyes by Phanerochaete chrysosporium and Pleurotus sajorcaju. *Enzyme and Microbial Technology,* 29, 473-477.

CHAIRATTANAMANOKORN, P., IMAI, T., KONDO, R., UKITA, M. & PRASERTSAN, P. (2006) Screening thermotolerant white-rot fungi for decolorization of wastewaters. *Applied Biochemistry and Biotechnology,* 128, 195-204.

CHEFETZ, B., CHEN, Y. & HADAR, Y. (1998) Purification and characterization of laccase from Chaetomium thermophilium and its role in humification. *Applied and Environmental Microbiology,* 64, 3175-3179.

CHOMCZYNSKI, P. & SACCHI, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate phenol chloroform extraction. *Analytical Biochemistry,* 162, 156-159.

CLAUS, H. (2004) Laccases: Structure, reactions, distribution. *Micron,* 35, 93-96.

COLAO, M. C., LUPINO, S., GARZILLO, A. M., BUONOCORE, V. & RUZZI, M. (2006) Heterologous expression of lccl gene from Trametes trogii in Pichia pastoris and characterization of the recombinant enzyme. *Microbial Cell Factories,* 5, -.

COUTO, S. R. (2007) Laccase production at reactor scale by filamentous fungi *Biotechnology Advances,* 25, 558-569.

DASHTBAN, M., SCHRAFT, H. & QIN, W. (2009) Fungal bioconversion of lignocellulosic residues; Opportunities & perspectives. *International Journal of Biological Sciences,* 5, 578-595.

DAVIS, B. J. (1964) Disc electrophoresis 2. Method and application to human serum proteins. *Annals of the New York Academy of Sciences,* 121, 404-&.

DIAZ, E. (2007) *Microbial biodegradation: genomics and molecular biology,* Caister Academic Press.

DITTMER, J. K., PATEL, N. J., DHAWALE, S. W. & DHAWALE, S. S. (1997) Production of multiple laccase isoforms by Phanerochaete chrysosporium grown under nutrient sufficiency. *Fems Microbiology Letters,* 149, 65-70.

DOI, R. H. & KOSUGI, A. (2004) Cellulosomes: Plant-cell-wall-degrading enzyme complexes. *Nature Reviews Microbiology,* 2, 541-551.

DON, R. H., COX, P. T., WAINWRIGHT, B. J., BAKER, K. & MATTICK, J. S. (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research,* 19, 4008.

DUCROS, V., BRZOZOWSKI, A. M., WILSON, K. S., ØSTERGAARD, P., SCHNEIDER, P., A.;, S. & DAVIES, G. J. (2001) Structure of the laccase from Coprinus cinereus at 1.68 Å resolution: evidence for different `type 2 Cu-depleted' isoforms. *Acta Cryst.,* D57.

DÜMMLER, A., LAWRENCE, A.-M. & DE MARCO, A. (2005) Simplified screening for the detection of soluble fusion constructs expressed in E. coli using a modular set of vectors. *Microbial Cell Factories,* 4.

EDENS, W. A., GOINS, T. Q., DOOLEY, D. & HENSON, J. M. (1999) Purification and characterization of a secreted laccase of Gaeumannomyces graminis var. tritici. *Applied and Environmental Microbiology,* 65, 3071-3074.

EISENBERG, D., LUTHY, R. & BOWIE, J. U. (1997) VERIFY3D: Assessment of protein models with three-dimensional profiles. *Macromolecular Crystallography, partt B,* 277 396-404.

ERKURT, H. A. (2010) Biodegradation of Azo Dyes. *The handbook of environmental chemistry.* Springer VBH.

FARHANGRAZI, Z. S., SINCLAIR, R., YAMAZAKI, I. & POWERS, L. S. (1992) Haloperoxidase Activity of Phanerochaete-Chrysosporium Lignin Peroxidase-H2 and Peroxidase-H8. *Biochemistry,* 31, 10763-10768.

FERRARONI, M., MYASOEDOVA, N. M., SCHMATCHENKO, V., LEONTIEVSKY, A. A., GOLOVLEVA, L. A., SCOZZAFAVA, A. & BRIGANTI, F. (2007) Crystal structure of a blue laccase from Lentinus tigrinus: evidences for intermediates in the molecular oxygen reductive splitting by multicopper oxidases. *Bmc Structural Biology,* 7, -.

FINN, R. D., TATE, J., MISTRY, J., COGGILL, P. C., SAMMUT, S. J., HOTZ, H.-R., CERIC, G., FORSLUND, K., EDDY, S. R., SONNHAMMER, E. L. L. & BATEMAN, B. (2008) The Pfam protein families database. *Nucl. Acids Res.* , 36 (suppl 1), D281-D288.

FLURKEY, W. H., BETTY RATCLIFF, B., LOPEZ, L., KUGLIN, J. & DAWLEY, R. M. (1995) Differentiation of Fungal Tyrosinases and Laccases Using Selective Inhibitors and Substrates IN LEE, Y. C. & WHITAKER, J. R. (Eds.) *Enzymatic Browning and Its Prevention.*

FRASCONI, M., FAVERO, G., BOER, H., KOIVULA, A. & MAZZEI, F. (2010) Kinetic and biochemical properties of high and low redox potential laccases from fungal and plant origin. *Biochimica et Biophysica Acta - Proteins and Proteomics,* 1804, 899-908.

GALAGAN, J. E., CALVO, S. E., CUOMO, C., MA, L.-J., WORTMAN, J. R., BATZOGLOU, S., LEE, S.-I., BASTURKMEN, M., SPEVAK, C. C., CLUTTERBUCK, J., KAPITONOV, V., JURKA, J., SCAZZOCCHIO, C., FARMAN, M., BUTLER, J., PURCELL, S., HARRIS, S., BRAUS, G. H., DRAHT, O., BUSCH, S., D'ENFERT, C., BOUCHIER, C., GOLDMAN, G. H., BELL-PEDERSEN, D., GRIFFITHS-JONES, S., DOONAN, J. H., YU, J., VIENKEN, K., PAIN, A., FREITAG, M., SELKER, E. U., ARCHER, D. B., PENALVA, M. A., OAKLEY, B. R., MOMANY, M., TANAKA, T., KUMAGAI, T., ASAI, K., MACHIDA, M., NIERMAN, W. C., DENNING, D. W., CADDICK, M., HYNES, M., PAOLETTI, M., FISCHER, R., MILLER, B., DYER, P., SACHS, M. S., OSMANI, S. A. & BIRREN, B. W. (2005) Sequencing of Aspergillus nidulans and comparative analysis with A. fumigatus and A. oryzae. *Nature,* 438, 1105-1115.

GARZILLO, A. M. V., COLAO, M. C., CARUSO, C., CAPORALE, C., CELLETTI, D. & BUONOCORE, V. (1998) Laccase from the white-rot fungus Trametes trogii. *Applied Microbiology and Biotechnology,* 49, 545-551.

GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S., WILKINS, M. R., APPEL, R. D. & BAIROCH, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. IN WALKER, J. M. (Ed.) *The Proteomics Protocols Handbook.* Humana Press.

GIARDINA, P., FARACO, V., PEZZELLA, C., PISCITELLI, A., VANHULLE, S. & SANNIA, G. (2010) Laccases: A never-ending story. *Cellular and Molecular Life Sciences,* 67, 369-385.

GIETZ, R. D. & SCHIESTL, R. H. (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols,* 2, 31-34.

GUEX, N. & PEITSCH, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis,* 18, 2714-2723.

GUPTA, R., JUNG, E. & BRUNAK, S. (2004) Prediction of N-glycosylation sites in human proteins.

HAGER, L. P., MORRIS, D. R., BROWN, F. S. & EBERWEIN, H. (1966) Chloroperoxidase. II. Utilization of halogen anions. *J Biol Chem,* 241, 1769-77.

HARDEN, A. (2009) *Alcoholic fermentation.*

HILDEN, K., HAKALA, T. K. & LUNDELL, T. (2009) Thermotolerant and thermostable laccases. *Biotechnology Letters,* 31, 1117-1128.

HIRD, H., PUMPHREY, R., WILSON, P., SUNDERLAND, J. & REECE, P. (2000) Identification of peanut and hazelnut allergens by native two-dimensional gel electrophoresis. *Electrophoresis,* 21, 2678-2683.

HOU, C. T. & SHAW, J.-F. (2009) *Biocatalysis and Agricultural Biotechnology*, CRC Press.

IFFLAND, A., GENDREIZIG, S., TAFELMEYER, P. & JOHNSSON, K. (2001) Changing the substrate specificity of cytochrome c peroxidase using directed evolution. *Biochemical and Biophysical Research Communications,* 286, 126-132.

JORDAAN, J. & LEUKES, W. D. (2003) Isolation of a thermostable laccase with DMAB and MBTH oxidative coupling activity from a mesophilic white rot fungus. *Enzyme and Microbial Technology,* 33, 212-219.

JUHASZ, A. L. & NAIDU, R. (2000) Bioremediation of high molecular weight polycyclic aromatic hydrocarbons: A review of the microbial degradation of benzo[a]pyrene. *International Biodeterioration and Biodegradation,* 45, 57-88.

KELLEY, L. A. & STERNBERG, M. J. E. (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nature Protocols,* 4.

KERSTEN, P. J., KALYANARAMAN, B., HAMMEL, K. E., REINHAMMAR, B. & KIRK, T. K. (1990) Comparison of lignin peroxidase, horseradish peroxidase and laccase in the oxidation of methoxybenzenes. *Biochemical Journal,* 268, 475-480.

KING, A. J., CRAGG, S. M., LI, Y., DYMOND, J., GUILLE, M. J., BOWLES, D. J., BRUCE, N. C., GRAHAM, I. A. & MCQUEEN-MASON, S. J. (2010) Molecular insight into lignocellulose digestion by a marine isopod in the absence of gut microbes. *Proceedings of the National Academy of Sciences,* 107, 5345-5350.

KIRBY, R. (2005) Actinomycetes and lignin degradation. *Advances in Applied Microbiology.*

KIRK, O., BORCHERT, T. V. & FUGLSANG, C. C. (2002) Industrial enzyme applications. *Current Opinion in Biotechnology,* 13, 345-351.

KIRK, T. K. & CULLEN, D. (1998) Enzymology and molecular genetics of wood degradation by white-rot fungi. *Environmentally Friendly Technologies for the Pulp and Paper Industry.* Wiley.

KNITTEL, T. & PICARD, D. (1993) PCR with degenerate primers containing deoxyinosine fails with Pfu DNA polymerase. *PCR Methods Appl,* 2, 346-7.

KOSINSKI, J., CYMERMAN, I. A., FEDER, M., KUROWSKI, M. A., SASIN, J. M. & BUJNICKI, J. M. (2003) A "Frankenstein's monster" approach to comparative modeling: Merging the finest fragments of fold-recognition models and iterative model refinement aided by 3D structure evaluation. *Proteins-Structure Function and Genetics,* 53, 369-379.

KOUA, D., CERUTTI, L., FALQUET, L., SIGRIST, C. J., THEILER, G., HULO, N. & DUNAND, C. (2008) PeroxiBase: a database with new tools for peroxidase family classification. *Nucleic Acids Res.*

KUNAMNENI, A., CAMARERO, S., GARCIA-BURGOS, C., PLOU, F. J., BALLESTEROS, A. & ALCALDE, M. (2008a) Engineering and Applications of fungal laccases for organic synthesis. *Microbial Cell Factories,* 7, -.

KUNAMNENI, A., GHAZI, I., CAMARERO, S., BALLESTEROS, A., PLOU, F. J. & ALCALDE, M. (2008b) Decolorization of synthetic dyes by laccase immobilized on epoxy-activated carriers. *Process Biochemistry,* 43, 169-178.

LAEMMLI, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227, 680-685.

LE NGOC HUYEN, T., RÉMOND, C., DHEILLY, R. M. & CHABBERT, B. (2010) Effect of harvesting date on the composition and saccharification of Miscanthus x giganteus. *Bioresource Technology,* 101, 8224-8231.

LEONOWICZ, A., MATUSZEWSKA, A., LUTEREK, J., ZIEGENHAGEN, D., WOJTAŁ›-WASILEWSKA, M., CHO, N. S., HOFRICHTER, M. & ROGALSKI, J. (1999) Biodegradation of lignin by white rot fungi. *Fungal Genetics and Biology,* 27, 175-185.

LEONTIEVSKY, A., MYASOEDOVA, N., POZDNYAKOVA, N. & GOLOVLEVA, L. (1997a) 'Yellow' laccase of Panus tigrinus oxidizes non-phenolic substrates without electron-transfer mediators. *Febs Letters,* 413, 446-448.

LEONTIEVSKY, A. A., POZDNYAKOVA, N. N. & MYASOEDOVA, N. M. (1996) Comparative characterization of oxidase-1 from Panus tigrinus 8-18 and laccase from Coriolus versicolor VKM F-116

Biochemistry-Moscow, 61, 1262-1268.

LEONTIEVSKY, A. A., VARES, T., LANKINEN, P., SHERGILL, J. K., POZDNYAKOVA, N. N., MYASOEDOVA, N. M., KALKKINEN, N., GOLOVLEVA, L. A., CAMMACK, R., THURSTON, C. F. & HATAKKA, A. (1997b) Blue and yellow laccases of ligninolytic fungi. *Fems Microbiology Letters,* 156, 9-14.

LISOVA, Z., LISOV, A. & LEONTIEVSKY, A. (2010) Two laccase isoforms of the basidiomycete Cerrena unicolor VKMF-3196. Induction, isolation and properties *Journal of Basic Microbiology,* 50, 72-82.

LITTLECHILD, J. (1999) Haloperoxidases and their role in biotransformation reactions. *Curr Opin Chem Biol,* 3, 28-34.

LOBOS, S., LARRAÍN, J., SALAS, L., CULLEN, D. & VICUÑA, R. (1994) Isoenzymes of manganese-dependent peroxidase and laccase produced by the lignin-degrading basidiomycete Ceriporiopsis subvermispora. *Microbiology* 140.

LOPEZ, C., MIELGO, I., MOREIRA, M. T., FEIJOO, G. & LEMA, J. M. (2002) Enzymatic membrane reactors for biodegradation of

recalcitrant compounds. Application to dye decolourisation. *Journal of Biotechnology,* 99, 249-257.
LUNDELL, T. K., MAKELA, M. R. & HILDEN, K. (2010) Lignin-modifying enzymes in filamentous basidiomycetes - Ecological, functional and phylogenetic review. *Journal of Basic Microbiology,* 50, 5-20.
LÜTHY, R., BOWIE, J. U. & EISENBERG, D. (1992) Assessment of protein models with three-dimensional profiles. *Nature,* 356, 83-5.
MANOJ, K. M. & HAGER, L. P. (2006) A colorimetric method for detection and quantification of chlorinating activity of hemeperoxidases. *Analytical Biochemistry,* 348, 84-86.
MARTINEZ, A. T. (2002) Molecular biology and structure-function of lignin- degrading heme peroxidases. *Enzyme and Microbial Technology,* 30, 425-444.
MARTINEZ, A. T., SPERANZA, M., RUIZ-DUENAS, F. J., FERREIRA, P., CAMARERO, S., GUILLAN, F., MARTINEZ, M. J., GUTIĂ©RREZ, A. & DEL RUO, J. C. (2005) Biodegradation of lignocellulosics: Microbial, chemical, and enzymatic aspects of the fungal attack of lignin. *International Microbiology,* 8, 195-204.
MARTINS, L. O., SOARES, C. M., PEREIRA, M. M., TEIXEIRA, M., COSTA, T., JONES, G. H. & HENRIQUES, A. O. (2002) Molecular and biochemical characterization of a highly stable bacterial laccase that occurs as a structural component of the Bacillus subtilis endospore coat. *Journal of Biological Chemistry,* 277, 18849-18859.
MARZULLO, L., CANNIO, R., GIARDINA, P., SANTINI, M. T. & SANNIA, G. (1995) Veratryl alcohol oxidase from Pleurotus ostreatus participates in lignin biodegradation and prevents polymerization of laccase-oxidized substrates. *Journal of Biological Chemistry,* 270, 3823-3827.
MAY, S. W. (1999) Applications of oxidoreductases. *Current Opinion in Biotechnology,* 10, 370-375.
METZKER, M. L. (2010) Sequencing technologies the next generation. *Nature Reviews Genetics,* 11, 31-46.
MEYER, F., PAARMANN, D., D'SOUZA, M., OLSON, R., GLASS, E. M., KUBAL, M., PACZIAN, T., RODRIGUEZ, A., STEVENS, R., WILKE, A., WILKENING, J. & EDWARDS, R. A. (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics,* 9, 386.
MICHNIEWICZ, A., LEDAKOWICZ, S., ULLRICH, R. & HOFRICHTER, M. (2008) Kinetics of the enzymatic decolorization of textile dyes by laccase from Cerrena unicolor. *Dyes and Pigments,* 77, 295-302.
MIELE, A., GIARDINA, P., SANNIA, G. & FARACO, V. (2009) Random mutants of a Pleurotus ostreatus laccase as new

biocatalysts for industrial effluents bioremediation. *Journal of Applied Microbiology,* 108, 998-1006.
MIN, K. L., KIM, Y. H., YOUNG WOON, K., HACK SUNG, J. & YUNG CHIL, H. (2001) Characterization of a novel laccase produced by the wood-rotting fungus Phellinus ribis. *Archives of Biochemistry and Biophysics,* 392, 279-286.
MOGENSEN, A. S., DOLFING, J., HAAGENSEN, F. & AHRING, B. K. (2003) Potential for anaerobic conversion of xenobiotics. *Advances in biochemical engineering/biotechnology,* 82, 69-134.
MOLDES, D. (2004) Different proportions of laccase isoenzymes produced by submerged cultures of Trametes versicolor grown on lignocellulosic wastes. *Biotechnology Letters,* 26, 327-330.
MOORE, J. K. (2007) Pursuit of novel haloperoxidase for assymetric biocatalysis. *School of Biological Sciences.* Liverpool, University of Liverpool.
MORAWSKY, B., QUAN, S. & ARNOLD, F. H. (2001) Functional expression and stabilization of horseradish peroxidase by directed evolution in Sacharomyces cerevisiae. *Biotechnology and Bioengineering* 76, 101–107.
MOROZOVA, O. V., SHUMAKOVICH, G. P., GORBACHEVA, M. A., SHLEEV, S. V. & YAROPOLOV, A. I. (2007) "Blue" laccases. *Biochemistry (Moscow),* 72, 1136-1150.
MORRIS, D. R. & HAGER, L. P. (1966) Chloroperoxidase. I. Isolation and properties of the crystalline glycoprotein. *J Biol Chem,* 241, 1763-8.
NAKAMURA, K. & GO, N. (2005) Function and molecular evolution of multicopper blue proteins. *Cellular and Molecular Life Sciences,* 62, 2050-2066.
NIELSEN, H., ENGELBRECHT, J., BRUNAK, S. & VON HEIJNE, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng. 1997 Jan;10(1):1-6.,* 10, 1-6.
NOVY, R. & MORRIS, B. (2001) Use of glucose to control basal expression in the pET System. *inNovations.*
NYAHONGO, G. S., SCHROEDER, M., STEINER, W. & GÄL'BITZ, G. M. (2005) Biodegradation of 2,4,6-trinitrotoluene (TNT): An enzymatic perspective. *Biocatalysis and Biotransformation,* 23, 53-69.
OCHMAN, H., GERBER, A. S. & HARTL, D. L. (1988) Genetic Applications of an Inverse Polymerase Chain Reaction. *Genetics,* 120, 621-623.
OHTSUKA, E., MATSUKI, S., IKEHARA, M., TAKAHASHI, Y. & MATSUBARA, K. (1985) An alternative approach to deoxyoligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions. *Journal of Biological Chemistry,* 260, 2605-2608.

ORNSTEIN, L. (1964) Disc electrophoresis 1. Background and theory. *Annals of the New York Academy of Sciences,* 121, 321-&.

ORTIZ-BERMUDEZ, P., HIRTH, K. C., SREBOTNIK, E. & HAMMEL, K. E. (2007) Chlorination of lignin by ubiquitous fungi has a likely role in global organochlorine production. *Proceedings of the National Academy of Sciences of the United States of America,* 104, 3895-3900.

PALMIERI, G., GIARDINA, P., BIANCO, C., SCALONI, A., CAPASSO, A. & SANNIA, G. (1997) A novel white laccase from Pleurotus ostreatus. *Journal of Biological Chemistry,* 272, 31301-31307.

PANDEY, A., SINGH, P. & IYENGAR, L. (2007) Bacterial decolorization and degradation of azo dyes. *International Biodeterioration and Biodegradation,* 59, 73-84.

PATEL, R. N. (2006) *Biocatalysis in the pharmaceutical and biotechnology industries*, CRC Press.

PATRICK, C. H. (1952) *Alcohol, Culture, and Society*, Duke University Press.

PECYNA, M. J., ULLRICH, R., BITTNER, B., CLEMENS, A., SCHEIBNER, K., SCHUBERT, R. & HOFRICHTER, M. (2009) Molecular characterization of aromatic peroxygenase from Agrocybe aegerita. *Applied Microbiology and Biotechnology,* 84, 885-897.

PENG, R. H., XIONG, A. S., XUE, Y., FU, X. Y., GAO, F., ZHAO, W., TIAN, Y. S. & YAO, Q. H. (2008) Microbial biodegradation of polyaromatic hydrocarbons. *FEMS Microbiology Reviews,* 32, 927-955.

PEREIRA, L., VIEGAS, C., COELHO, A., ROBALO, M. P. & MARTINS, L. (2007) Mechanisms of enzymatic degradation of azo and anthraquinone dyes by bacterial CotA-laccase. *Journal of Biotechnology,* 131, S107-S108.

PEREZ, J., MUNOZ-DORADO, J., DE LA RUBIA, T. & MARTINEZ, J. (2002) Biodegradation and biological treatments of cellulose, hemicellulose and lignin: An overview. *International Microbiology,* 5, 53-63.

PIONTEK, K., ANTORINI, M. & CHOINOWSKI, T. (2002) Crystal structure of a laccase from the fungus Trametes versicolor at 1.90-A resolution containing a full complement of coppers. *Journal of Biological Chemistry,* 277, 37663-37669.

PIONTEK, K., ULLRICH, R., LIERS, C., DIEDERICHS, K., PLATTNER, D. A. & HOFRICHTER, M. (2010) Crystallization of a 45 kDa peroxygenase/peroxidase from the mushroom Agrocybe aegerita and structure determination by SAD utilizing only the haem iron. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications,* 66, 693-698.

POZDNYAKOVA, N. N., RODAKIEWICZ-NOWAK, J. & TURKOVSKAYA, O. V. (2004) Catalytic properties of yellow laccase from Pleurotus ostreatus D1. *Journal of Molecular Catalysis B-Enzymatic,* 30, 19-24.

POZDNYAKOVA, N. N., RODAKIEWICZ-NOWAK, J., TURKOVSKAYA, O. V. & HABER, J. (2006) Oxidative degradation of polyaromatic hydrocarbons and their derivatives catalyzed directly by the yellow laccase from Pleurotus ostreatus D1. *Journal of Molecular Catalysis B-Enzymatic,* 41, 8-15.

REED, K. C. & MANN, D. A. (1985) Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Research,* 13, 7207-7221.

REGALADO, C. (2004) Biotechnological applications of peroxidases. *Phytochemistry reviews,* 3 243 -256

RIVA, S. (2006) Laccases: blue enzymes for green chemistry. *Trends in Biotechnology,* 24, 219-226.

ROBERTS, M. B. V. (1986) *Biology - A Functional Approach Fourth Edition.*

RODGERS, C. J., BLANFORD, C. F., GIDDENS, S. R., SKAMNIOTI, P., ARMSTRONG, F. A. & GURR, S. J. (2009) Designer laccases: a vogue for high-potential fungal enzymes? *Trends in Biotechnology,* 28, 63-72.

RODRIGUEZ, E., RUIZ-DUENAS, F. J., KOOISTRA, R., RAM, A., MARTINEZ, A. T. & MARTINEZ, M. J. (2008) Isolation of two laccase genes from the white-rot fungus Pleurotus eryngii and heterologous expression of the pel3 encoded protein. *Journal of Biotechnology,* 134, 9-19.

ROZZELL, J. D. (1999) Commercial scale biocatalysis: Myths and realities. *Bioorganic and Medicinal Chemistry,* 7, 2253-2261.

RYDBERG, P., SIGFRIDSSON, E. & RYDE, U. (2004) On the role of the axial ligand in heme proteins: a theoretical study *Journal of Biological Inorganc Chemistry,* 9, 203-223.

SAKAKIBARA, A. (1977) Degradation products of protolignin and the structure of lignin. In: The Structure, Biosynthesis, and Degradation of Wood. IN LOEWUS, F. A. & RUNECKLES, V. C. (Eds.) *Recent Advances in Phytochemistry, Vol. 11.* New York, Plenum Press.

SALI, A. & BLUNDELL, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* , 234, 779-815.

SAMBROOK J., F. E. F. A. M. T. (1989) Molecular Cloning - A Laboratory Manual.

SANCHEZ, C. (2009) Lignocellulosic residues: Biodegradation and bioconversion by fungi. *Biotechnology Advances,* 27, 185-194.

SCHOEMAKER, H. E., MINK, D. & WUBBOLTS, M. G. (2003) Dispelling the myths - Biocatalysis in industrial synthesis. *Science,* 299, 1694-1697.

SCHULZE, B. & WUBBOLTS, M. G. (1999) Biocatalysis for industrial production of fine chemicals. *Current Opinion in Biotechnology,* 10, 609-615.

SELINHEIMO, E., KRUUS, K., BUCHERT, J., HOPIA, A. & AUTIO, K. (2006) Effects of laccase, xylanase and their combination on the rheological properties of wheat doughs. *Journal of Cereal Science,* 43, 152-159.

SHENG, D. & GOLD, M. H. (1997) Haloperoxidase activity of manganese peroxidase from Phanerochaete chrysosporium. *Arch Biochem Biophys,* 345, 126-34.

SHIA, J., SHARMA-SHIVAPPA, R. R. & CHINNA, M. S. (2009) Microbial pretreatment of cotton stalks by submerged cultivation of Phanerochaete chrysosporium. *Bioresource Technology,* 100 4388-4395

SHORT, J. M., FERNANDEZ, J. M., SORGE, J. A. & HUSE, W. D. (1988) Lambda-ZAP - a bacteriophage lambda-expression vector with in vivo excision properties. *Nucleic Acids Research,* 16, 7583-7600.

SINGH, A. & WARD, O. P. (2004) *Biodegradation and bioremediation,* Springer.

SINHA, S., CHATTOPADHYAY, P., PAN, I., CHATTERJEE, S., CHANDA, P., BANDYOPADHYAY, D., DAS, K. & SEN, S. K. (2009) Microbial transformation of xenobiotics for environmental bioremediation. *African Journal of Biotechnology,* 8, 6016-6027.

SOUTHERN, E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology,* 98, 503-517.

STRAATHOF, A. J. J. & ADLERCREUTZ, P. (2000) *Applied biocatalysis,* CRC Press.

SUNDARAMOORTHY, M., TERNER, J. & POULOS, T. L. (1995) The crystal structure of chloroperoxidase: a heme peroxidase--cytochrome P450 functional hybrid. *Structure,* 3, 1367-77.

TELLEZ-JURADO, A., ARANA-CUENCA, A., BECERRA, A. E. G., VINIEGRA-GONZALEZ, G. & LOERA, O. (2006) Expression of a heterologous laccase by Aspergillus niger cultured by solid-state and submerged fermentations. *Enzyme and Microbial Technology,* 38, 665-669.

ULLRICH, R., DOLGE, C., KLUGE, M. & HOFRICHTER, M. (2008) Pyridine as novel substrate for regioselective oxygenation with aromatic peroxygenase from Agrocybe aegerita. *Febs Letters,* 582, 4100-4106.

ULLRICH, R. & HOFRICHTER, M. (2005) The haloperoxidase of the agaric fungus Agrocybe aegerita hydroxylates toluene and naphthalene. *FEBS Lett,* 579, 6247-50.

ULLRICH, R., NUSKE, J., SCHEIBNER, K., SPANTZEL, J. & HOFRICHTER, M. (2004) Novel haloperoxidase from the agaric

basidiomycete Agrocybe aegerita oxidizes aryl alcohols and aldehydes. *Appl Environ Microbiol,* 70, 4575-81.
VAN PEE, K. H., DONG, C., FLECKS, S., NAISMITH, J., PATALLO, E. P. & WAGE, T. (2006) Biological halogenation has moved far beyond haloperoxidases. *Adv Appl Microbiol,* 59, 127-57.
VERA, J. C., WHEAT, C. W., FESCEMYER, H. W., FRILANDER, M. J., CRAWFORD, D. L., HANSKI, I. & MARDEN, J. H. (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Molecular Ecology,* 17, 1636-1647.
VERHAGEN, F. J. M., SWARTS, H. J., KUYPER, T. W., WIJNBERG, J. B. P. A. & FIELD, J. A. (1996) The ubiquity of natural adsorbable organic halogen production among basidiomycetes. *Applied Microbiology and Biotechnology,* 45, 710-718.
VOKOUNOVA, M., VACEK, O. & KUNC, F. (1992) Degradation of the herbicide bromoxynil in Pseudomonas putida. *Folia Microbiologica,* 37, 122-127.
VRIEND, G. (1990) WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.,* 8, 52-56.
WACKETT, L. P. & HERSHBERGER, C. D. (2001) *Biocatalysis and biodegradation*, ASM Press.
WAGNER, C., MOLITOR, I. M. & KONIG, G. M. (2008) Critical view on the monochlorodimedone assay utilized to detect haloperoxidase activity. *Phytochemistry,* 69, 323-32.
WARNECKE, F., LUGINBUHL, P. & IVANOVA, N. (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature,* 450, 560-565.
WELINDER, K. G. (1992) Superfamili of plant, fungal and bacterial peroxidases. *Current Opinion in Structural Biology,* 2, 388-393.
WELLS, A., TERIA, M. & EVE, T. (2006) Green oxidations with laccase-mediator systems. *Biochemical Society Transactions,* 34, 304-308.
WHITTALL, J. & SUTTON, P. (2009) *Practical Methods for Biocatalysis and Biotransformations*, Wiley-Blackwell
WIDSTEN, P. & KANDELBAUER, A. (2008) Laccase applications in the forest products industry: A review. *Enzyme and Microbial Technology,* 42, 293-307.
YAMASAKI, C., KAWASHIMA, H., TODOKORO, F., IMAMIZU, Y., OGAWA, M., TANINO, M., ITOH, T., GOJOBORI, T. & IMANISHI, T. (2006) TACT: Transcriptome Auto-annotation Conducting Tool of H-InvDB. *Nucleic Acids Research,* 34, W345-W349.
YOSHIDA, H. (1883) Chemistry of lacquer (Urushi). Part 1. *J. Chem.Soc.,* 43, 472-486.
YOSHIDA, M., LIU, Y., UCHIDA, S., KAWARADA, K., UKAGAMI, Y., ICHINOSE, H., KANEKO, S. & FUKUDA, K. (2008) Effects of cellulose crystallinity, hemicellulose, and lignin on the

enzymatic hydrolysis of Miscanthus sinensis to monosaccharides. *Bioscience Biotechnology and Biochemistry,* 72, 805-810.

YOUNGS, H. L., MOENNE-LOCCOZ, P., LOEHR, T. M. & GOLD, M. H. (2000) Formation of a Bis(histidyl) Heme Iron Complex in Manganese Peroxidase at High pH and Restoration of the Native Enzyme Structure by Calcium. *Biochemistry,* 39, 9994-10000.

ZAKS, A. (2001) Industrial biocatalysis. *Current Opinion in Chemical Biology,* 5, 130-136.

ZHAO, H., CHOCKALINGAM, K. & CHEN, Z. (2002) Directed evolution of enzymes and pathways for industrial biocatalysis. *Current Opinion in Biotechnology,* 13, 104-110.

ZHAO, H. & VAN DER DONK, W. A. (2003) Regeneration of cofactors for use in biocatalysis. *Current Opinion in Biotechnology,* 14, 583-589.

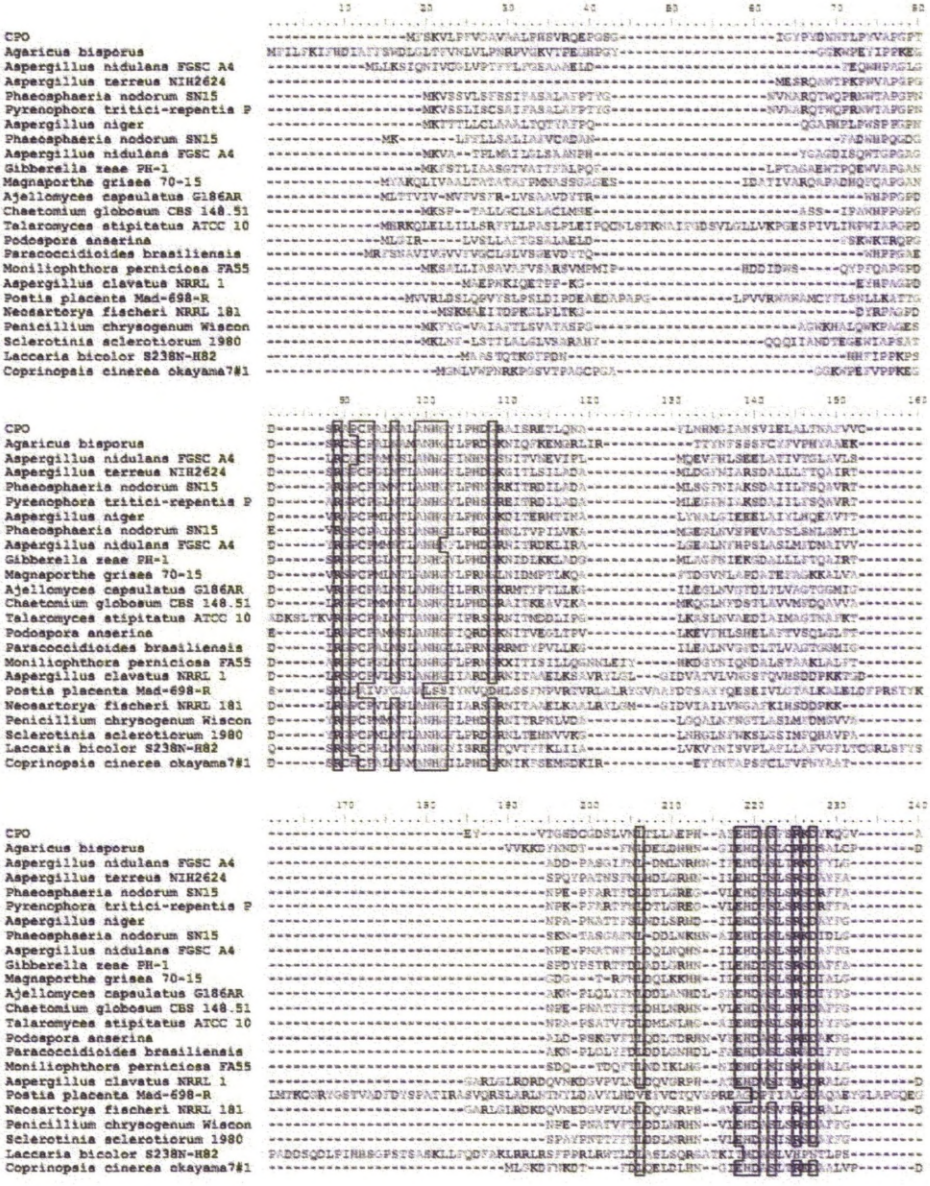
ZUMARRAGA, M., CAMARERO, S., SHLEEV, S., MARTINEZ-ARIAS, A., BALLESTEROS, A., PLOU, F. J. & ALCALDE, M. (2008) Altering the laccase functionality by in vivo assembly of mutant libraries with different mutational spectra. *Proteins-Structure Function and Bioinformatics,* 71, 250-260.

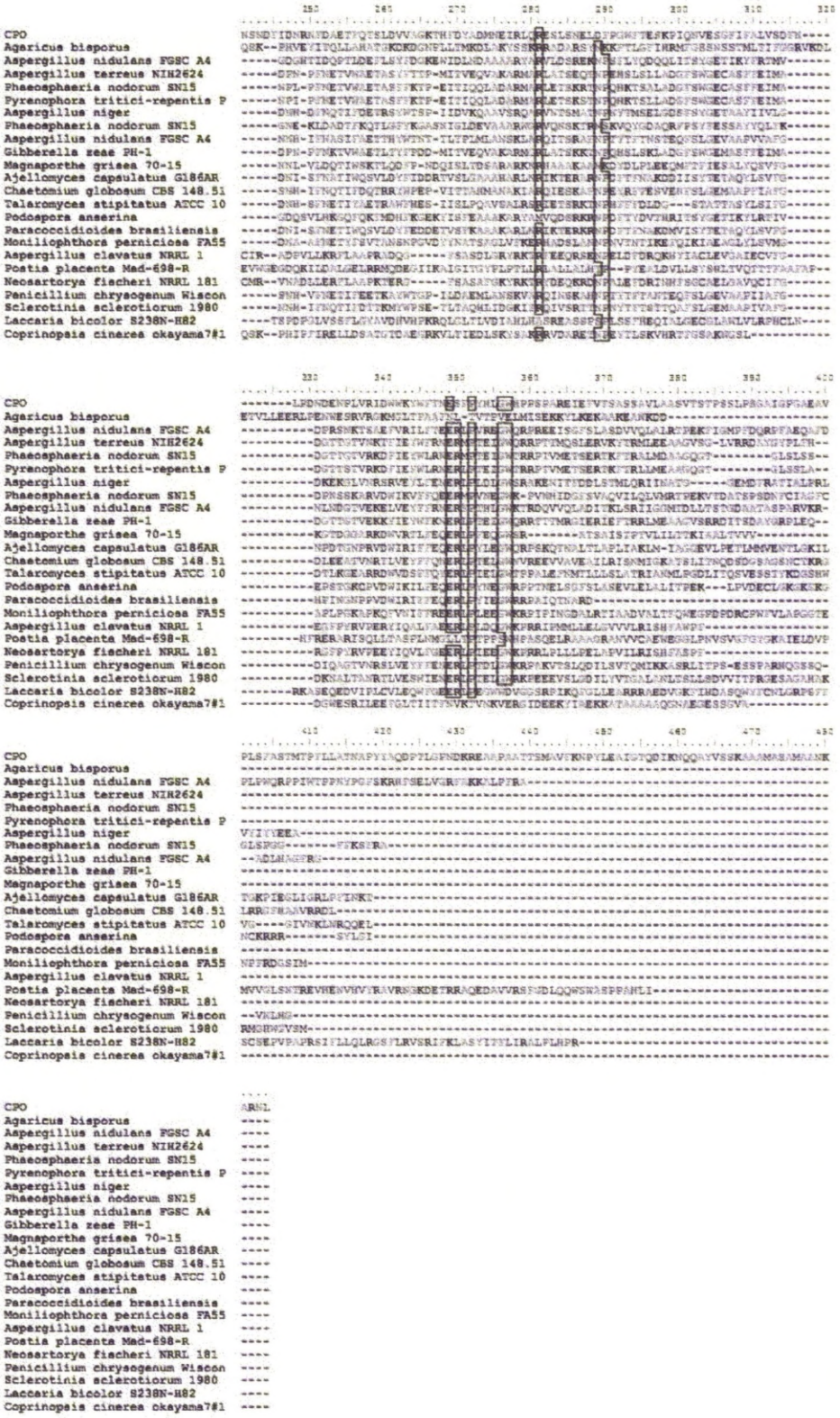
Appendices

SUPPLEMENTARY DATA

1. Appendix 1 – Alignments and sequence data.

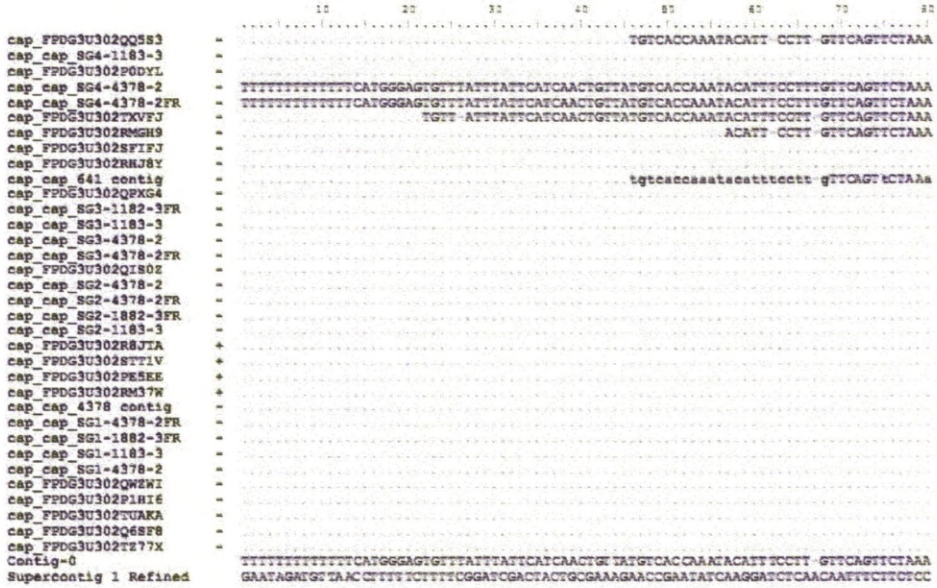
1.1. CPO protein alignment with AN7823 and other related proteins.

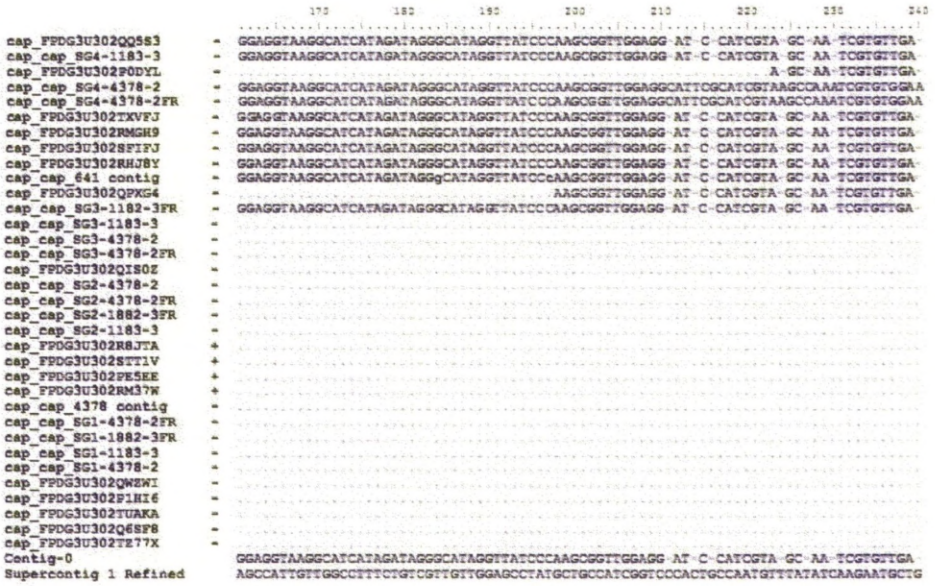


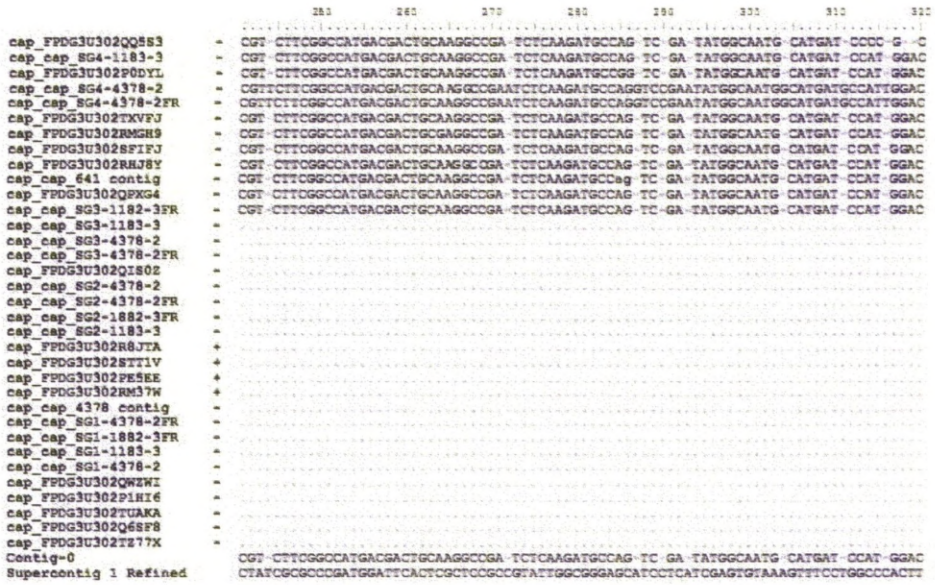

CPO
Agaricus bisporus
Aspergillus nidulans FGSC A4
Aspergillus terreus NIH2624
Phaeosphaeria nodorum SN15
Pyrenophora tritici-repentis P
Aspergillus niger
Phaeosphaeria nodorum SN15
Aspergillus nidulans FGSC A4
Gibberella zeae PH-1
Magnaporthe grisea 70-15
Ajellomyces capsulatus G186AR
Chaetomium globosum CBS 148.51
Talaromyces stipitatus ATCC 10
Podospora anserina
Paracoccidioides brasiliensis
Moniliophthora perniciosa FA55
Aspergillus clavatus NRRL 1
Postia placenta Mad-698-R
Neosartorya fischeri NRRL 181
Penicillium chrysogenum Wisconsin
Sclerotinia sclerotiorum 1980
Laccaria bicolor S238N-H82
Coprinopsis cinerea okayama7#1

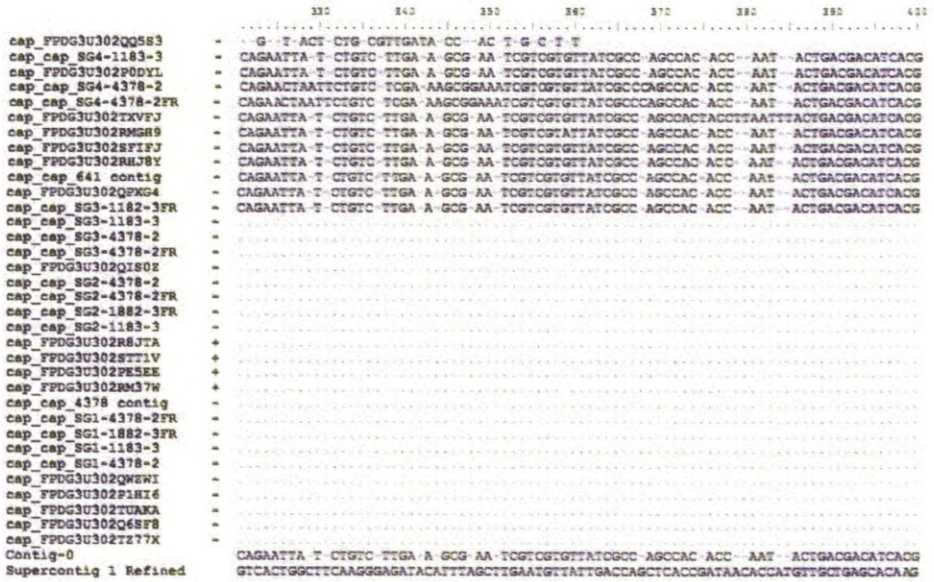
1.2. Assembly of *S. aeruginosa* laccase maxi- and super- contigs.

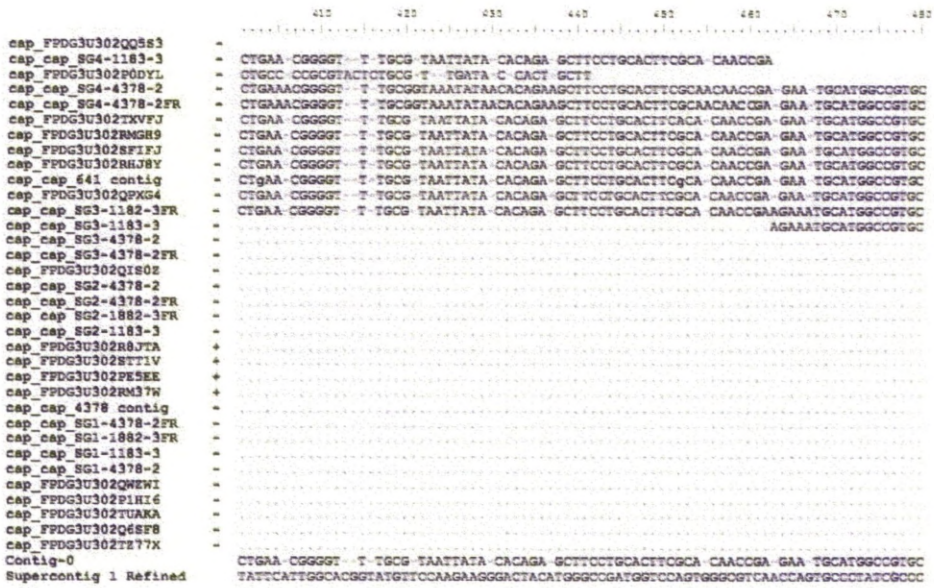
Assembly of supercontig 1 corresponding to the laccase gene N2C1.

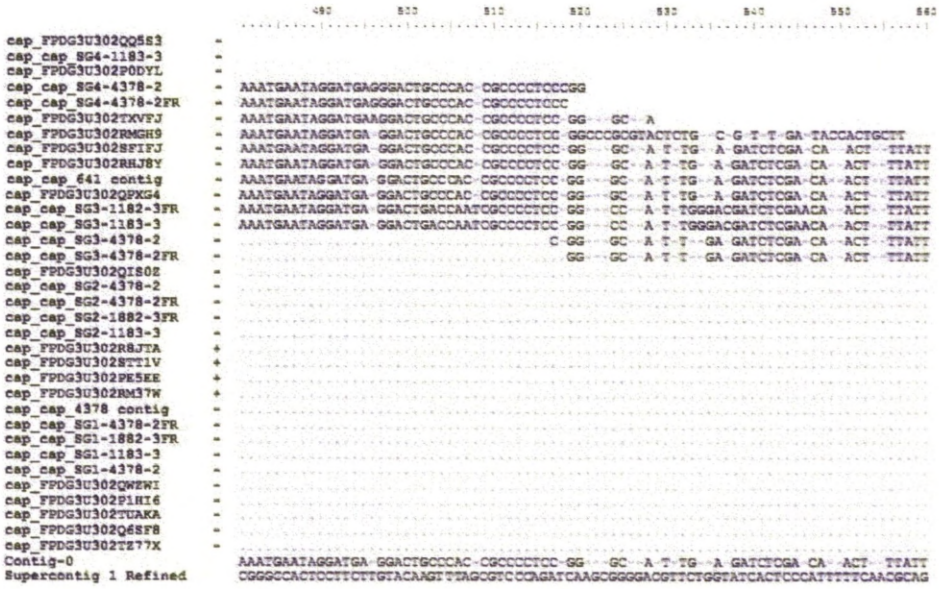


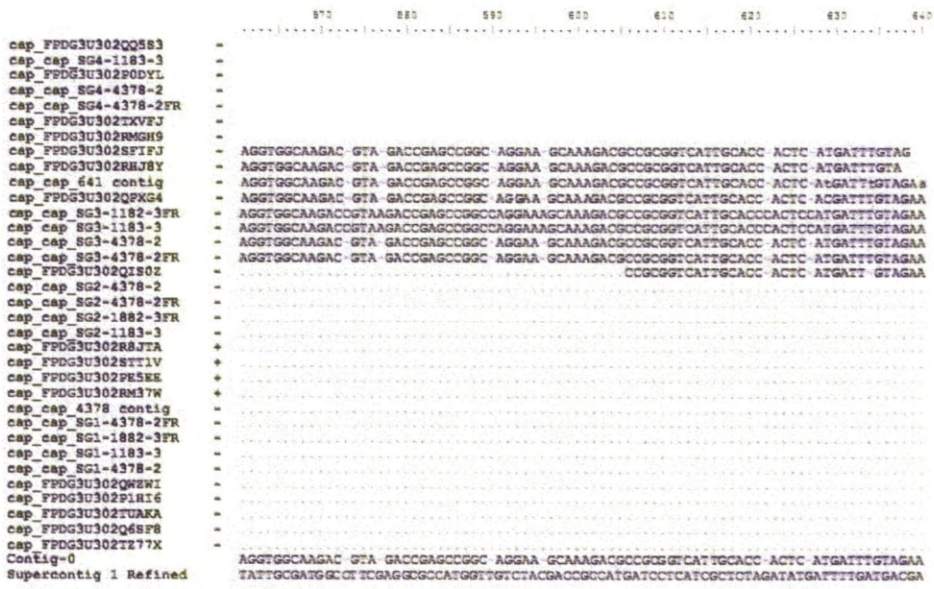

170 180 190 200 210 220 230 240
cap_FPDG3U302QQ5S3 - GGAGGTAAGGCATCATAGATAGGGCATAGGTTATCCCAAGCGGTTGGAGG AT C CATCGTA GC AA TCGTGTTGA
cap_cap_SG4-1183-3 - GGAGGTAAGGCATCATAGATAGGGCATAGGTTATCCCAAGCGGTTGGAGG AT C CATCGTA GC AA TCGTGTTGA
cap_FPDG3U302P0DYL - A GC AA TCGTGTTGA
cap_cap_SG4-4378-2 - GGAGGTAAGGCATCATAGATAGGGCATAGGTTATCCCAAGCGGTTGGAGGCATTCGCATCGTAAGCCAAATCGTGTGGAA
cap_cap_SG4-4378-2FR - GGAGGTAAGGCATCATAGATAGGGCATAGGTTATCCCAAGCGGTTGGAGGCATTCGCATCGTAAGCCAAATCGTGTGGAA
cap_FPDG3U302TXVFJ - GGAGGTAAGGCATCATAGATAGGGCATAGGTTATCCCAAGCGGTTGGAGG AT C CATCGTA GC AA TCGTGTTGA
cap_FPDG3U302RMGH9 - GGAGGTAAGGCATCATAGATAGGGCATAGGTTATCCCAAGCGGTTGGAGG AT C CATCGTA GC AA TCGTGTTGA
cap_FPDG3U302SFIFJ - GGAGGTAAGGCATCATAGATAGGGCATAGGTTATCCCAAGCGGTTGGAGG AT C CATCGTA GC AA TCGTGTTGA
cap_FPDG3U302RHJ8Y - GGAGGTAAGGCATCATAGATAGGGCATAGGTTATCCCAAGCGGTTGGAGG AT C CATCGTA GC AA TCGTGTTGA
cap_cap_641 contig - GGAGGTAAGGCATCATAGATAGGgCATAGGTTATCCcAAGCGGTTGGAGG AT C CATCGTA GC AA TCGTGTTGA
cap_FPDG3U302QPXG4 - AAGCGGTTGGAGG AT C CATCGTA GC AA TCGTGTTGA
cap_cap_SG3-1182-3FR - GGAGGTAAGGCATCATAGATAGGGCATAGGTTATCCCAAGCGGTTGGAGG AT C CATCGTA GC AA TCGTGTTGA
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z -
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR -
cap_cap_SG2-1183-3 -
cap_FPDG3U302R8JTA +
cap_FPDG3U302STT1V +
cap_FPDG3U302PE5EE +
cap_FPDG3U302RM37W +
cap_cap_4378 contig -
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI -
cap_FPDG3U302P1HI6 -
cap_FPDG3U302TUAKA -
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302TZ77X -
Contig-0 GGAGGTAAGGCATCATAGATAGGGCATAGGTTATCCCAAGCGGTTGGAGG AT C CATCGTA GC AA TCGTGTTGA
Supercontig 1 Refined AGCCATTGTTGGCCTTTCTGTCGTTGTTGGAGCCTATGCTGCCATCGGTCCCACTGCCAATGTTTATATCAAGAATGCTG

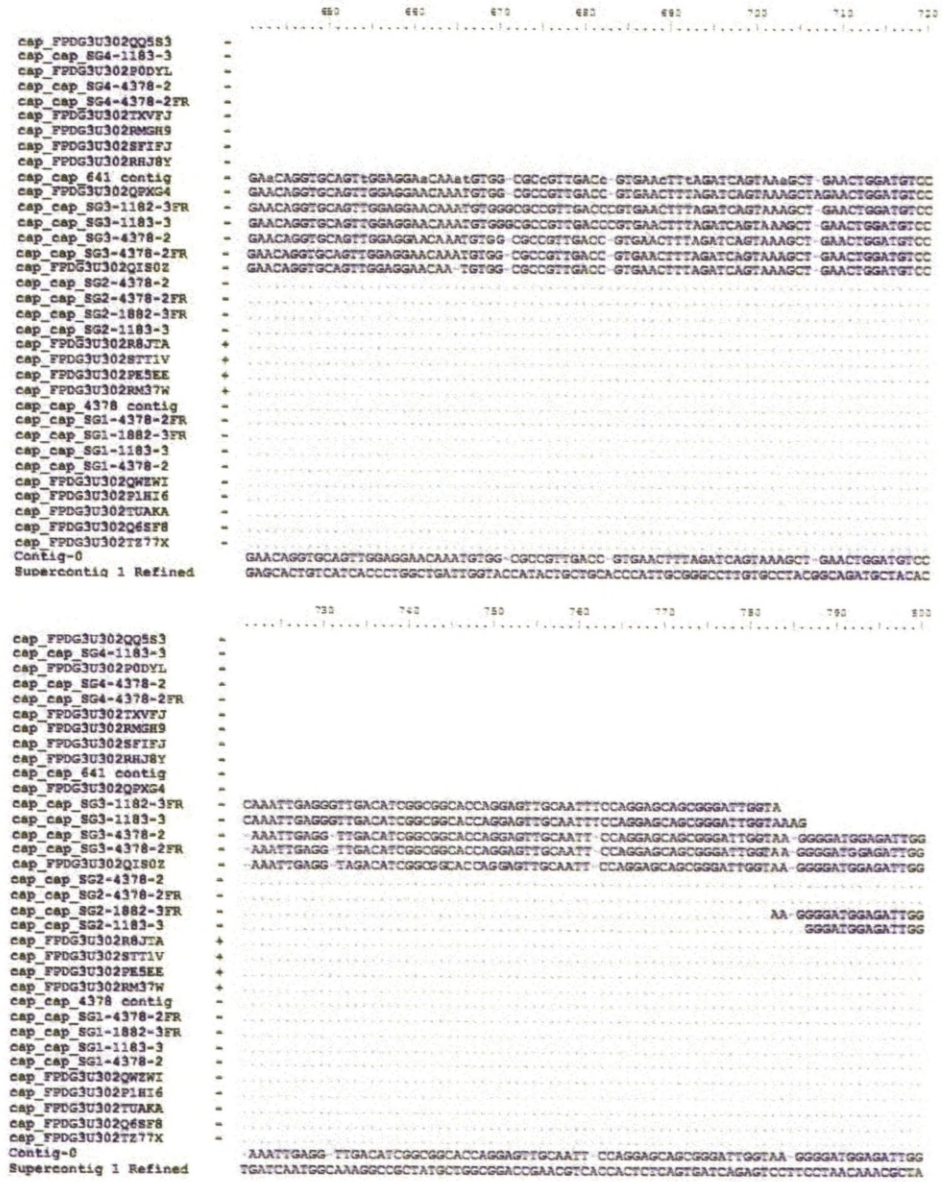

250 260 270 280 290 300 310 320
cap_FPDG3U302QQ5S3 - CGT CTTCGGCCATGACGACTGCAAGGCCGA TCTCAAGATGCCAG TC GA TATGGCAATG CATGAT CCCC G C
cap_cap_SG4-1183-3 - CGT CTTCGGCCATGACGACTGCAAGGCCGA TCTCAAGATGCCAG TC GA TATGGCAATG CATGAT CCAT GGAC
cap_FPDG3U302P0DYL - CGT CTTCGGCCATGACGACTGCAAGGCCGA TCTCAAGATGCCGG TC GA TATGGCAATG CATGAT CCAT GGAC
cap_cap_SG4-4378-2 - CGTTCTTCGGCCATGACGACTGCAAGGCCGAATCTCAAGATGCCAGGTCCGAATATGGCAATGGCATGATGCCATTGGAC
cap_cap_SG4-4378-2FR - CGTCTTCGGCCATGACGACTGCAAGGCCGAATCTCAAGATGCCAGGTCCGAATATGGCAATGGCATGATGCCATTGGAC
cap_FPDG3U302TXVFJ - CGT CTTCGGCCATGACGACTGCAAGGCCGA TCTCAAGATGCCAG TC GA TATGGCAATG CATGAT CCAT GGAC
cap_FPDG3U302RMGH9 - CGT CTTCGGCCATGACGACTGCGAGGCCGA TCTCAAGATGCCAG TC GA TATGGCAATG CATGAT CCAT GGAC
cap_FPDG3U302SFIFJ - CGT CTTCGGCCATGACGACTGCAAGGCCGA TCTCAAGATGCCAG TC GA TATGGCAATG CATGAT CCAT GGAC
cap_FPDG3U302RHJ8Y - CGT CTTCGGCCATGACGACTGCAAGGCGA TCTCAAGATGCCAG TC GA TATGGCAATG CATGAT CCAT GGAC
cap_cap_641 contig - CGT CTTCGGCCATGACGACTGCAAGGCCGA TCTCAAGATGCCag TC GA TATGGCAATG CATGAT CCAT GGAC
cap_FPDG3U302QPXG4 - CGT CTTCGGCCATGACGACTGCAAGGCCGA TCTCAAGATGCCAG TC GA TATGGCAATG CATGAT CCAT GGAC
cap_cap_SG3-1182-3FR - CGT CTTCGGCCATGACGACTGCAAGGCCGA TCTCAAGATGCCAG TC GA TATGGCAATG CATGAT CCAT GGAC
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z -
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR -
cap_cap_SG2-1183-3 -
cap_FPDG3U302R8JTA +
cap_FPDG3U302STT1V +
cap_FPDG3U302PE5EE +
cap_FPDG3U302RM37W +
cap_cap_4378 contig -
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI -
cap_FPDG3U302P1HI6 -
cap_FPDG3U302TUAKA -
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302TZ77X -
Contig-0 CGT CTTCGGCCATGACGACTGCAAGGCCGA TCTCAAGATGCCAG TC GA TATGGCAATG CATGAT CCAT GGAC
Supercontig 1 Refined CTATCGCGCCCGATGGATTCACTCGCTCCGCCGTATTGGCGGGAGCATCCTCATCGAGTGTAAAGTTTCCTGGCCCACTT

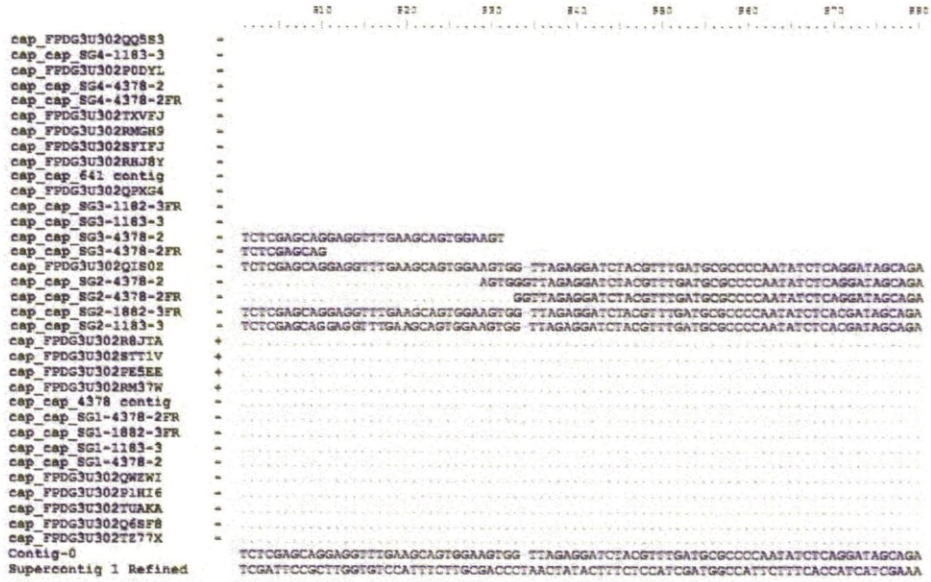

cap_FPDG3U302QQ5S3
cap_cap_SG4-1183-3
cap_FPDG3U302P0DYL
cap_cap_SG4-4378-2
cap_cap_SG4-4378-2FR
cap_FPDG3U302TXVFJ
cap_FPDG3U302RMGH9
cap_FPDG3U302SFIFJ
cap_FPDG3U302RHJ8Y
cap_cap_641 contig
cap_FPDG3U302QPXG4
cap_cap_SG3-1182-3FR
cap_cap_SG3-1183-3
cap_cap_SG3-4378-2
cap_cap_SG3-4378-2FR
cap_FPDG3U302QIS0Z
cap_cap_SG2-4378-2
cap_cap_SG2-4378-2FR
cap_cap_SG2-1882-3FR
cap_cap_SG2-1183-3
cap_FPDG3U302R8JTA
cap_FPDG3U302STT1V
cap_FPDG3U302PE5EE
cap_FPDG3U302RM37W
cap_cap_4378 contig
cap_cap_SG1-4378-2FR
cap_cap_SG1-1882-3FR
cap_cap_SG1-1183-3
cap_cap_SG1-4378-2
cap_FPDG3U302QWEWI
cap_FPDG3U302P1HI6
cap_FPDG3U302TUAKA
cap_FPDG3U302Q6SF8
cap_FPDG3U302TZ77X
Contig-0
Supercontig 1 Refined

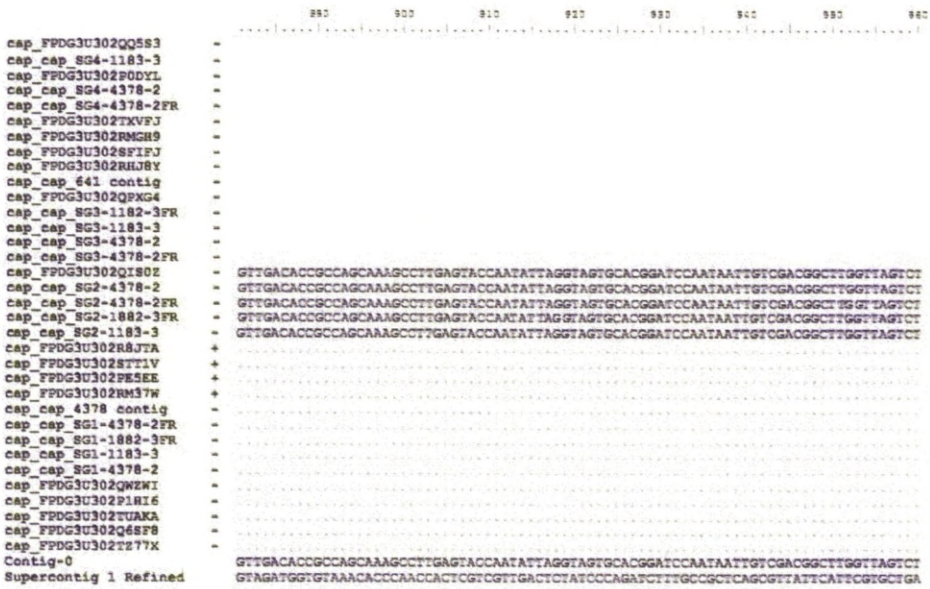

cap_FPDG3U302QQ5S3
cap_cap_SG4-1183-3
cap_FPDG3U302P0DYL
cap_cap_SG4-4378-2
cap_cap_SG4-4378-2FR
cap_FPDG3U302TXVFJ
cap_FPDG3U302RMGH9
cap_FPDG3U302SFIFJ
cap_FPDG3U302RHJ8Y
cap_cap_641 contig
cap_FPDG3U302QPXG4
cap_cap_SG3-1182-3FR
cap_cap_SG3-1183-3
cap_cap_SG3-4378-2
cap_cap_SG3-4378-2FR
cap_FPDG3U302QIS0Z
cap_cap_SG2-4378-2
cap_cap_SG2-4378-2FR
cap_cap_SG2-1882-3FR
cap_cap_SG2-1183-3
cap_FPDG3U302R8JTA
cap_FPDG3U302STT1V
cap_FPDG3U302PE5EE
cap_FPDG3U302RM37W
cap_cap_4378 contig
cap_cap_SG1-4378-2FR
cap_cap_SG1-1882-3FR
cap_cap_SG1-1183-3
cap_cap_SG1-4378-2
cap_FPDG3U302QWEWI
cap_FPDG3U302P1HI6
cap_FPDG3U302TUAKA
cap_FPDG3U302Q6SF8
cap_FPDG3U302TZ77X
Contig-0
Supercontig 1 Refined

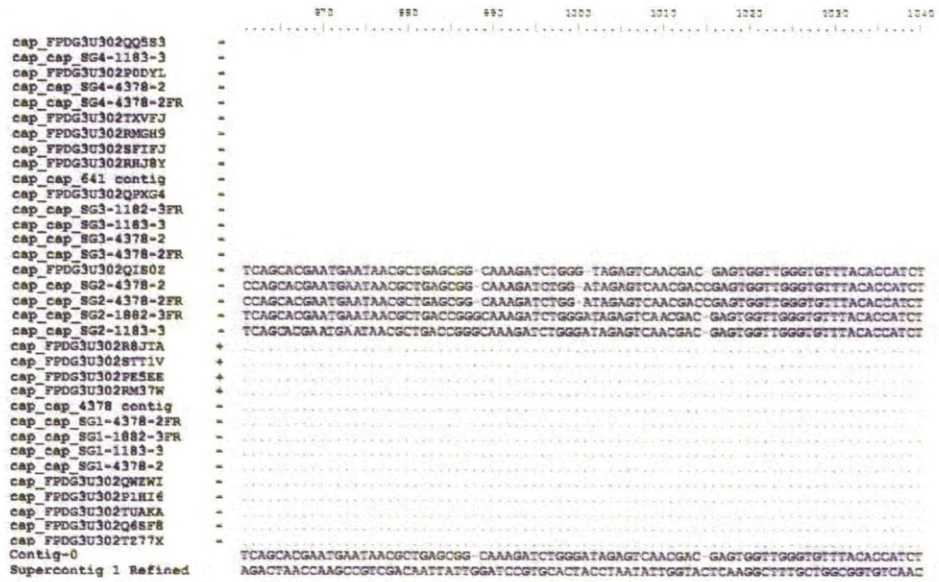

cap_FPDG3U302QQ5S3
cap_cap_SG4-1183-3
cap_FPDG3U302P0DYL
cap_cap_SG4-4378-2
cap_cap_SG4-4378-2FR
cap_FPDG3U302TXVFJ
cap_FPDG3U302RMGH9
cap_FPDG3U302SFIFJ
cap_FPDG3U302RHJ8Y
cap_cap_641 contig
cap_FPDG3U302QPXG4
cap_cap_SG3-1182-3FR
cap_cap_SG3-1183-3
cap_cap_SG3-4378-2
cap_cap_SG3-4378-2FR
cap_FPDG3U302QIS0Z
cap_cap_SG2-4378-2
cap_cap_SG2-4378-2FR
cap_cap_SG2-1882-3FR
cap_cap_SG2-1183-3
cap_FPDG3U302R8JTA
cap_FPDG3U302STT1V
cap_FPDG3U302PE5EE
cap_FPDG3U302RM37W
cap_cap_4378 contig
cap_cap_SG1-4378-2FR
cap_cap_SG1-1882-3FR
cap_cap_SG1-1183-3
cap_cap_SG1-4378-2
cap_FPDG3U302QWZWI
cap_FPDG3U302P1HI6
cap_FPDG3U302TUAKA
cap_FPDG3U302Q6SF8
cap_FPDG3U302TZ77X
Contig-0
Supercontig 1 Refined

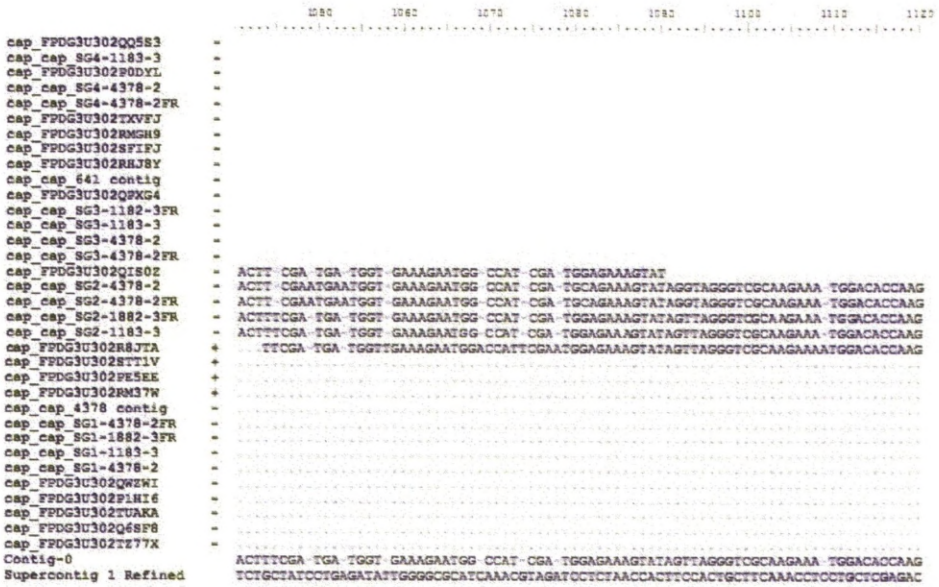

cap_FPDG3U302QQ5S3
cap_cap_SG4-1183-3
cap_FPDG3U302P0DYL
cap_cap_SG4-4378-2
cap_cap_SG4-4378-2FR
cap_FPDG3U302TXVFJ
cap_FPDG3U302RMGH9
cap_FPDG3U302SFIFJ
cap_FPDG3U302RHJ8Y
cap_cap_641 contig
cap_FPDG3U302QPXG4
cap_cap_SG3-1182-3FR
cap_cap_SG3-1183-3
cap_cap_SG3-4378-2
cap_cap_SG3-4378-2FR
cap_FPDG3U302QIS0Z
cap_cap_SG2-4378-2
cap_cap_SG2-4378-2FR
cap_cap_SG2-1882-3FR
cap_cap_SG2-1183-3
cap_FPDG3U302R8JTA
cap_FPDG3U302STT1V
cap_FPDG3U302PE5EE
cap_FPDG3U302RM37W
cap_cap_4378 contig
cap_cap_SG1-4378-2FR
cap_cap_SG1-1882-3FR
cap_cap_SG1-1183-3
cap_cap_SG1-4378-2
cap_FPDG3U302QWZWI
cap_FPDG3U302P1HI6
cap_FPDG3U302TUAKA
cap_FPDG3U302Q6SF8
cap_FPDG3U302TZ77X
Contig-0
Supercontig 1 Refined

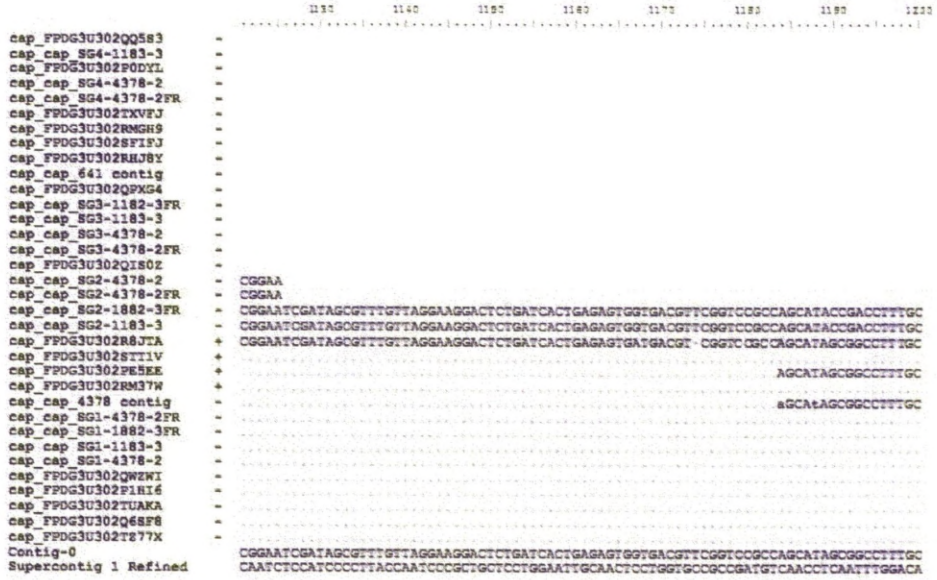

650 660 670 680 690 700 710 720
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig - GAaCAGGTGCAGTtGGAGGAaCAAatGTGG CGCcGTTGACc GTGAAcTTtAGATCAGTAAaGCT GAACTGGATGTCC
cap_FPDG3U302QPXG4 - GAACAGGTGCAGTTGGAGGAACAAATGTGG CGCCGTTGACC GTGAACTTTAGATCAGTAAAGCTAGAACTGGATGTCC
cap_cap_SG3-1182-3FR - GAACAGGTGCAGTTGGAGGAACAAATGTGGGCGCCGTTGACCCGTGAACTTTAGATCAGTAAAGCT GAACTGGATGTCC
cap_cap_SG3-1183-3 - GAACAGGTGCAGTTGGAGGAACAAATGTGGGCGCCGTTGACCCGTGAACTTTAGATCAGTAAAGCT GAACTGGATGTCC
cap_cap_SG3-4378-2 - GAACAGGTGCAGTTGGAGGAACAAATGTGG CGCCGTTGACC GTGAACTTTAGATCAGTAAAGCT GAACTGGATGTCC
cap_cap_SG3-4378-2FR - GAACAGGTGCAGTTGGAGGAACAAATGTGG CGCCGTTGACC GTGAACTTTAGATCAGTAAAGCT GAACTGGATGTCC
cap_FPDG3U302QIS0Z - GAACAGGTGCAGTTGGAGGAACAA TGTGG CGCCGTTGACC GTGAACTTTAGATCAGTAAAGCT GAACTGGATGTCC
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR -
cap_cap_SG2-1183-3 -
cap_FPDG3U302R8JTA +
cap_FPDG3U302STT1V +
cap_FPDG3U302PE5EE +
cap_FPDG3U302RM37W +
cap_cap_4378 contig -
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI -
cap_FPDG3U302P1HI6 -
cap_FPDG3U302TUAKA -
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302TZ77X -
Contig-0 GAACAGGTGCAGTTGGAGGAACAAATGTGG CGCCGTTGACC GTGAACTTTAGATCAGTAAAGCT GAACTGGATGTCC
Supercontig 1 Refined GAGCACTGTCATCACCCTGGCTGATTGGTACCATACTGCTGCACCCATTGCGGGCCTTGTGCCTACGGCAGATGCTACAC
730 740 750 760 770 780 790 800
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR - CAAATTGAGGGTTGACATCGGCGGCACCAGGAGTTGCAATTTCCAGGAGCAGCGGGATTGGTA
cap_cap_SG3-1183-3 - CAAATTGAGGGTTGACATCGGCGGCACCAGGAGTTGCAATTTCCAGGAGCAGCGGGATTGGTAAAG
cap_cap_SG3-4378-2 - AAATTGAGG TTGACATCGGCGGCACCAGGAGTTGCAATT CCAGGAGCAGCGGGATTGGTAA GGGGATGGAGATTGG
cap_cap_SG3-4378-2FR - AAATTGAGG TTGACATCGGCGGCACCAGGAGTTGCAATT CCAGGAGCAGCGGGATTGGTAA GGGGATGGAGATTGG
cap_FPDG3U302QIS0Z - AAATTGAGG TAGACATCGGCGGCACCAGGAGTTGCAATT CCAGGAGCAGCGGGATTGGTAA GGGGATGGAGATTGG
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR - AA GGGGATGGAGATTGG
cap_cap_SG2-1183-3 - GGGATGGAGATTGG
cap_FPDG3U302R8JTA +
cap_FPDG3U302STT1V +
cap_FPDG3U302PE5EE +
cap_FPDG3U302RM37W +
cap_cap_4378 contig -
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI -
cap_FPDG3U302P1HI6 -
cap_FPDG3U302TUAKA -
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302TZ77X -
Contig-0 AAATTGAGG TTGACATCGGCGGCACCAGGAGTTGCAATT CCAGGAGCAGCGGGATTGGTAA GGGGATGGAGATTGG
Supercontig 1 Refined TGATCAATGGCAAAGGCCGCTATGCTGGCGGACCGAACGTCACCACTCTCAGTGATCAGAGTCCTTCCTAACAAACGCTA

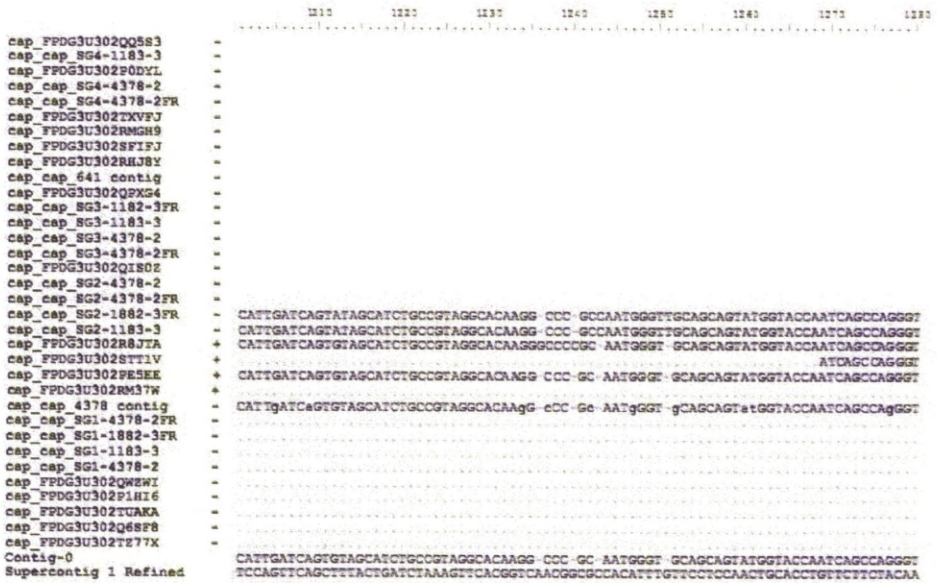

cap_FPDG3U302QQ5S3
cap_cap_SG4-1183-3
cap_FPDG3U302P0DYL
cap_cap_SG4-4378-2
cap_cap_SG4-4378-2FR
cap_FPDG3U302TXVFJ
cap_FPDG3U302RMGH9
cap_FPDG3U302SFIFJ
cap_FPDG3U302RHJ8Y
cap_cap_641 contig
cap_FPDG3U302QPXG4
cap_cap_SG3-1182-3FR
cap_cap_SG3-1183-3
cap_cap_SG3-4378-2
cap_cap_SG3-4378-2FR
cap_FPDG3U302QIS0Z
cap_cap_SG2-4378-2
cap_cap_SG2-4378-2FR
cap_cap_SG2-1882-3FR
cap_cap_SG2-1183-3
cap_FPDG3U302R8JTA
cap_FPDG3U302STT1V
cap_FPDG3U302PE5EE
cap_FPDG3U302RM37W
cap_cap_4378 contig
cap_cap_SG1-4378-2FR
cap_cap_SG1-1882-3FR
cap_cap_SG1-1183-3
cap_cap_SG1-4378-2
cap_FPDG3U302QWZWI
cap_FPDG3U302P1RI6
cap_FPDG3U302TUAKA
cap_FPDG3U302Q6SF8
cap_FPDG3U302TZ77X
Contig-0
Supercontig 1 Refined

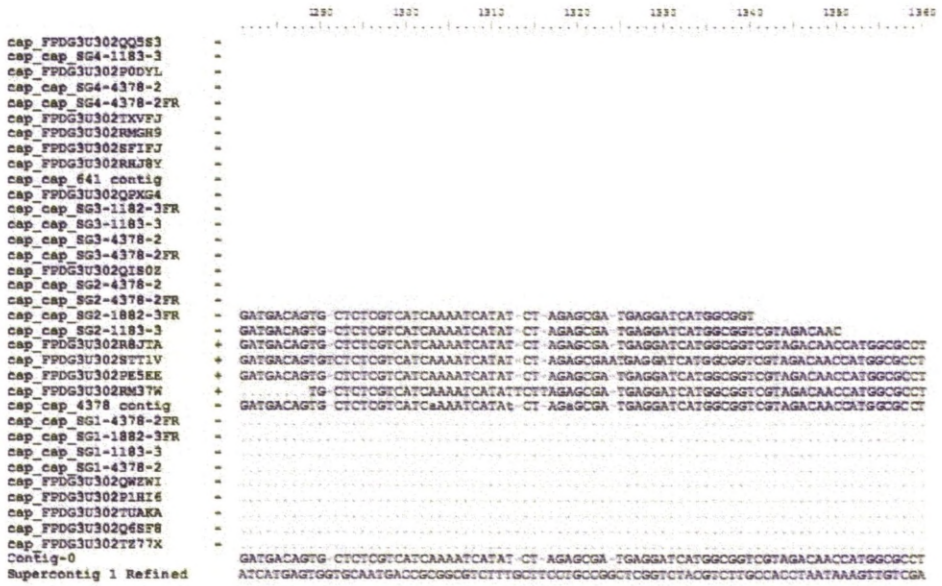

cap_FPDG3U302QQ5S3
cap_cap_SG4-1183-3
cap_FPDG3U302P0DYL
cap_cap_SG4-4378-2
cap_cap_SG4-4378-2FR
cap_FPDG3U302TXVFJ
cap_FPDG3U302RMGH9
cap_FPDG3U302SFIFJ
cap_FPDG3U302RHJ8Y
cap_cap_641 contig
cap_FPDG3U302QPXG4
cap_cap_SG3-1182-3FR
cap_cap_SG3-1183-3
cap_cap_SG3-4378-2
cap_cap_SG3-4378-2FR
cap_FPDG3U302QIS0Z
cap_cap_SG2-4378-2
cap_cap_SG2-4378-2FR
cap_cap_SG2-1882-3FR
cap_cap_SG2-1183-3
cap_FPDG3U302R8JTA
cap_FPDG3U302STT1V
cap_FPDG3U302PE5EE
cap_FPDG3U302RM37W
cap_cap_4378 contig
cap_cap_SG1-4378-2FR
cap_cap_SG1-1882-3FR
cap_cap_SG1-1183-3
cap_cap_SG1-4378-2
cap_FPDG3U302QWZWI
cap_FPDG3U302P1RI6
cap_FPDG3U302TUAKA
cap_FPDG3U302Q6SF8
cap_FPDG3U302TZ77X
Contig-0
Supercontig 1 Refined

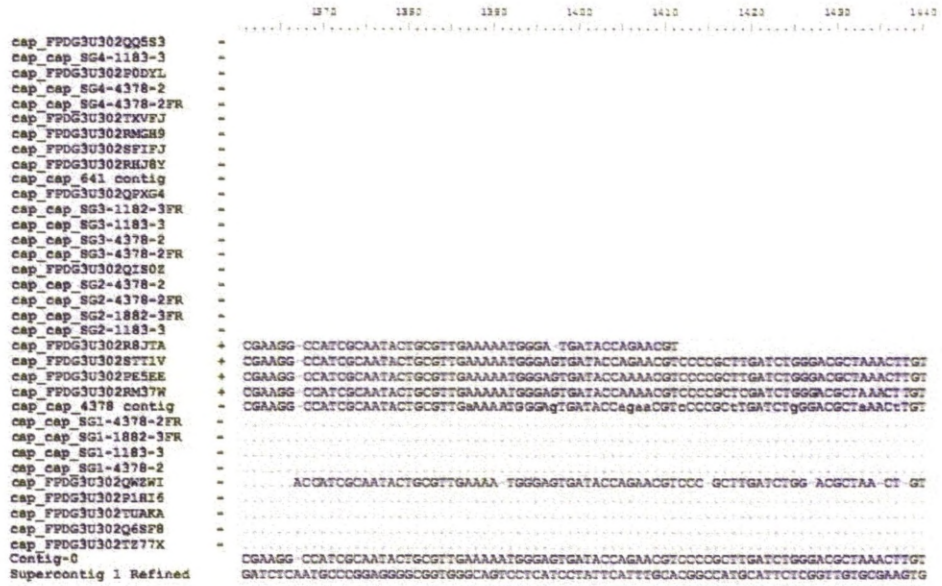

970 980 990 1000 1010 1020 1030 1040
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR -
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z - TCAGCACGAATGAATAACGCTGAGCGG CAAAGATCTGGG TAGAGTCAACGAC GAGTGGTTGGGTGTTTACACCATCT
cap_cap_SG2-4378-2 - CCAGCACGAATGAATAACGCTGAGCGG CAAAGATCTGG ATAGAGTCAACGACCGAGTGGTTGGGTGTTTACACCATCT
cap_cap_SG2-4378-2FR - CCAGCACGAATGAATAACGCTGAGCGG CAAAGATCTGG ATAGAGTCAACGACCGAGTGGTTGGGTGTTTACACCATCT
cap_cap_SG2-1882-3FR - TCAGCACGAATGAATAACGCTGACCGGGCAAAGATCTGGGATAGAGTCAACGAC GAGTGGTTGGGTGTTTACACCATCT
cap_cap_SG2-1183-3 - TCAGCACGAATGAATAACGCTGACCGGGCAAAGATCTGGGATAGAGTCAACGAC GAGTGGTTGGGTGTTTACACCATCT
cap_FPDG3U302R8JTA +
cap_FPDG3U302STT1V +
cap_FPDG3U302PE5EE +
cap_FPDG3U302RM37W +
cap_cap_4378 contig -
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI -
cap_FPDG3U302P1HI6 -
cap_FPDG3U302TUAKA -
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302TZ77X -
Contig-0 TCAGCACGAATGAATAACGCTGAGCGG CAAAGATCTGGGATAGAGTCAACGAC GAGTGGTTGGGTGTTTACACCATCT
Supercontig 1 Refined AGACTAACCAAGCCGTCGACAATTATTGGATCCGTGCACTACCTAATATTGGTACTCAAGGCTTTGCTGGCGGTGTCAAC

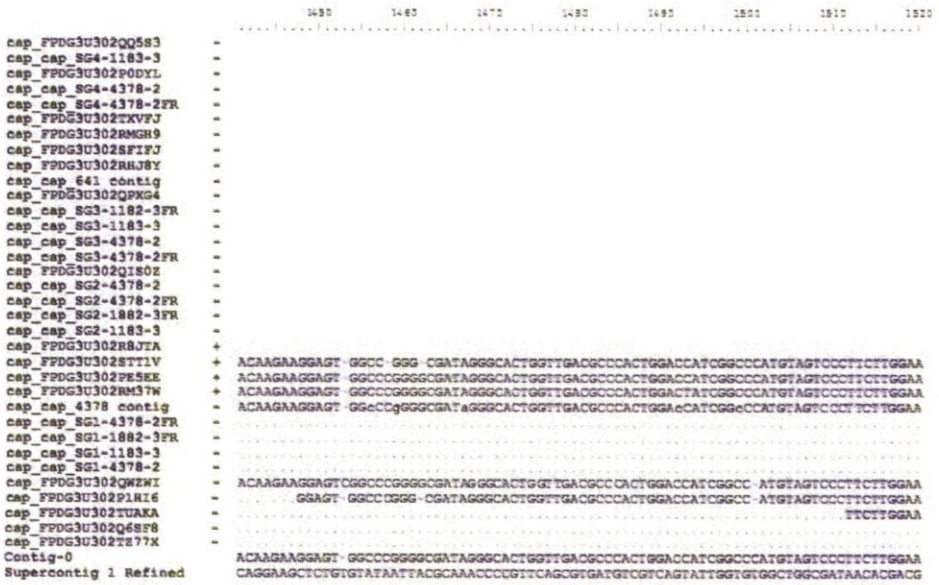

1080 1090 1070 1080 1090 1100 1110 1120
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR -
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z - ACTT CGA TGA TGGT GAAAGAATGG CCAT CGA TGGAGAAAGTAT
cap_cap_SG2-4378-2 - ACTT CGAATGAATGGT GAAAGAATGG CCAT CGA TGCAGAAAGTATAGGTAGGGTCGCAAGAAA TGGACACCAAG
cap_cap_SG2-4378-2FR - ACTT CGAATGAATGGT GAAAGAATGG CCAT CGA TGCAGAAAGTATAGGTAGGGTCGCAAGAAA TGGACACCAAG
cap_cap_SG2-1882-3FR - ACTTTCGA TGA TGGT GAAAGAATGG CCAT CGA TGGAGAAAGTATAGTTAGGGTCGCAAGAAA TGGACACCAAG
cap_cap_SG2-1183-3 - ACTTTCGA TGA TGGT GAAAGAATGG CCAT CGA TGGAGAAAGTATAGTTAGGGTCGCAAGAAA TGGACACCAAG
cap_FPDG3U302R8JTA + TTCGA TGA TGGTTGAAAGAATGGACCATTCGAATGGAGAAAGTATAGTTAGGGTCGCAAGAAAATGGACACCAAG
cap_FPDG3U302STT1V +
cap_FPDG3U302PE5EE +
cap_FPDG3U302RM37W +
cap_cap_4378 contig -
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI -
cap_FPDG3U302P1HI6 -
cap_FPDG3U302TUAKA -
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302TZ77X -
Contig-0 ACTTTCGA TGA TGGT GAAAGAATGG CCAT CGA TGGAGAAAGTATAGTTAGGGTCGCAAGAAA TGGACACCAAG
Supercontig 1 Refined TCTGCTATCCTGAGATATTGGGGCGCATCAAACGTAGATCCTCTAACCACTTCCACTGCTTCAAACCTCCTGCTCGAGAC

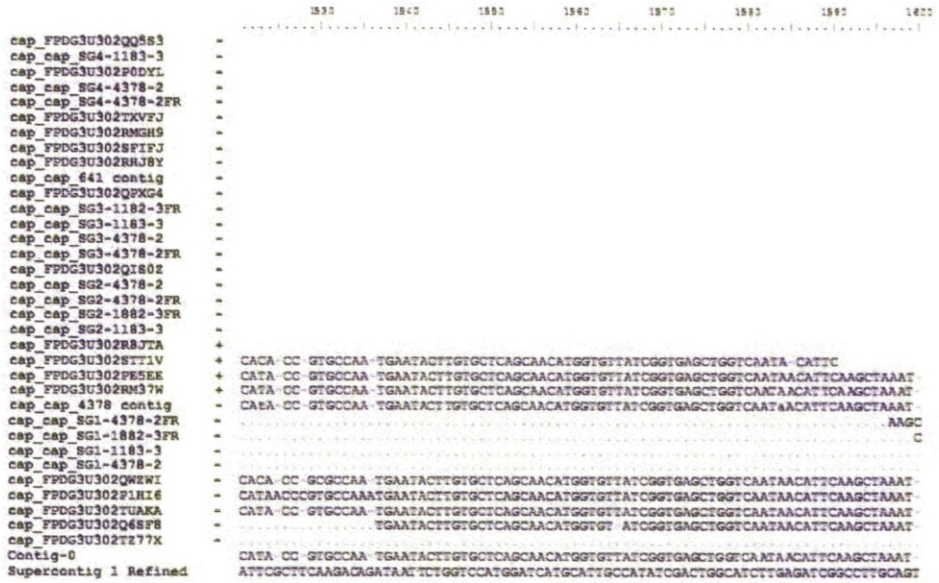

1130 1140 1150 1160 1170 1180 1190 1200
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR -
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z -
cap_cap_SG2-4378-2 - CGGAA
cap_cap_SG2-4378-2FR - CGGAA
cap_cap_SG2-1882-3FR - CGGAATCGATAGCGTTTGTTAGGAAGGACTCTGATCACTGAGAGTGGTGACGTTCGGTCCGCCAGCATACCGACCTTTGC
cap_cap_SG2-1183-3 - CGGAATCGATAGCGTTTGTTAGGAAGGACTCTGATCACTGAGAGTGGTGACGTTCGGTCCGCCAGCATACCGACCTTTGC
cap_FPDG3U302R8JTA + CGGAATCGATAGCGTTTGTTAGGAAGGACTCTGATCACTGAGAGTGATGACGT CGGTC CGC CAGCATAGCGGCCTTTGC
cap_FPDG3U302STT1V +
cap_FPDG3U302PE5EE + AGCATAGCGGCCTTTGC
cap_FPDG3U302RM37W +
cap_cap_4378 contig - aGCAtAGCGGCCTTTGC
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI -
cap_FPDG3U302P1HI6 -
cap_FPDG3U302TUAKA -
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302TZ77X -
Contig-0 CGGAATCGATAGCGTTTGTTAGGAAGGACTCTGATCACTGAGAGTGGTGACGTTCGGTCCGCCAGCATAGCGGCCTTTGC
Supercontig 1 Refined CAATCTCCATCCCCTTACCAATCCCGCTGCTCCTGGAATTGCAACTCCTGGTGCCGCCGATGTCAACCTCAATTTGGACA

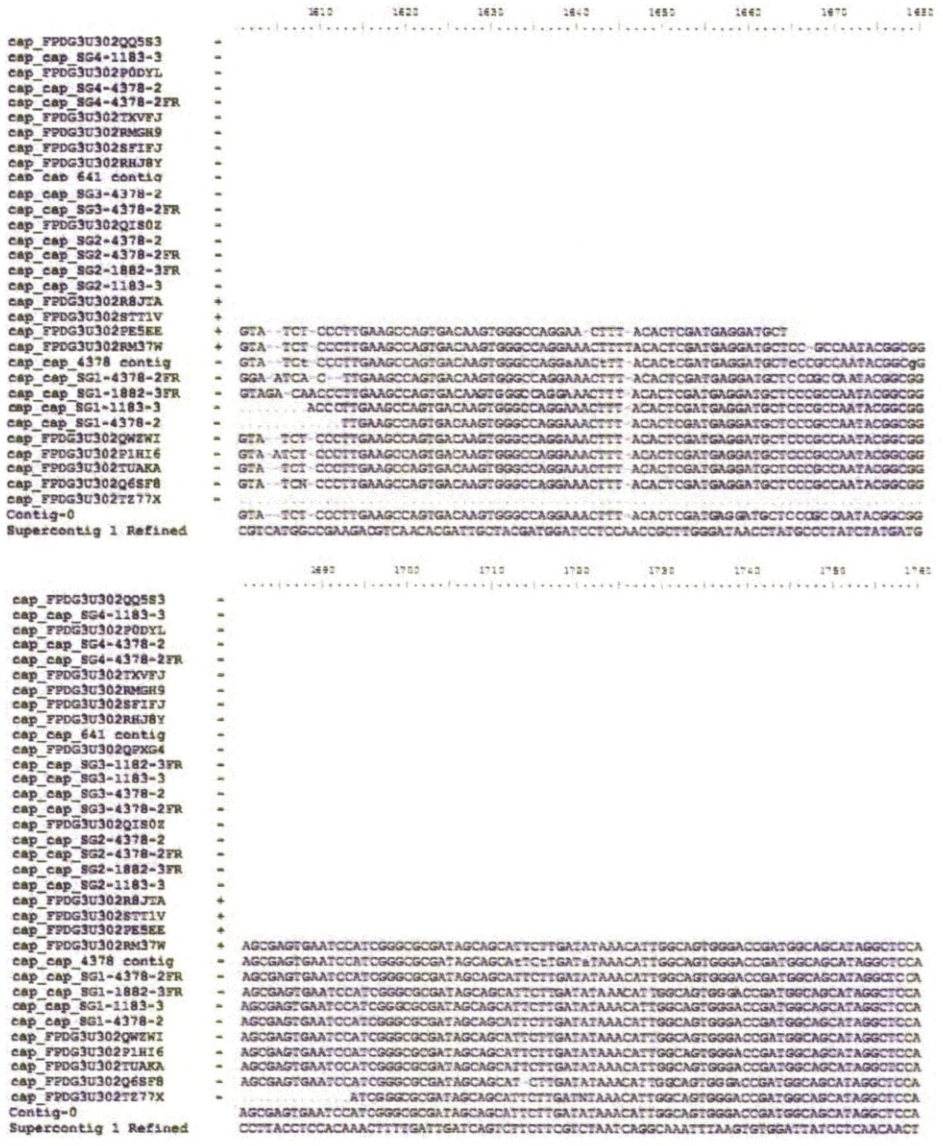

1210 1220 1230 1240 1250 1260 1270 1280
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR -
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z -
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR - CATTGATCAGTATAGCATCTGCCGTAGGCACAAGG CCC GCCAATGGGTTGCAGCAGTATGGTACCAATCAGCCAGGGT
cap_cap_SG2-1183-3 - CATTGATCAGTATAGCATCTGCCGTAGGCACAAGG CCC GCCAATGGGTTGCAGCAGTATGGTACCAATCAGCCAGGGT
cap_FPDG3U302R8JTA + CATTGATCAGTGTAGCATCTGCCGTAGGCACAAGGGCCCCGC AATGGGT GCAGCAGTATGGTACCAATCAGCCAGGGT
cap_FPDG3U302STT1V + ATCAGCCAGGGT
cap_FPDG3U302PE5EE + CATTGATCAGTGTAGCATCTGCCGTAGGCACAAGG CCC GC AATGGGT GCAGCAGTATGGTACCAATCAGCCAGGGT
cap_FPDG3U302RM37W +
cap_cap_4378 contig - CATTgATCaGTGTAGCATCTGCCGTAGGCACAAgG cCC Gc AATgGGT gCAGCAGTatGGTACCAATCAGCCAgGGT
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI -
cap_FPDG3U302P1HI6 -
cap_FPDG3U302TUAKA -
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302TZ77X -
Contig-0 CATTGATCAGTGTAGCATCTGCCGTAGGCACAAGG CCC GC AATGGGT GCAGCAGTATGGTACCAATCAGCCAGGGT
Supercontig 1 Refined TCCAGTTCAGCTTTACTGATCTAAAGTTCACGGTCAACGGCGCCACATTTGTTCCTCCAACTGCACCTGTTCTTCTACAA

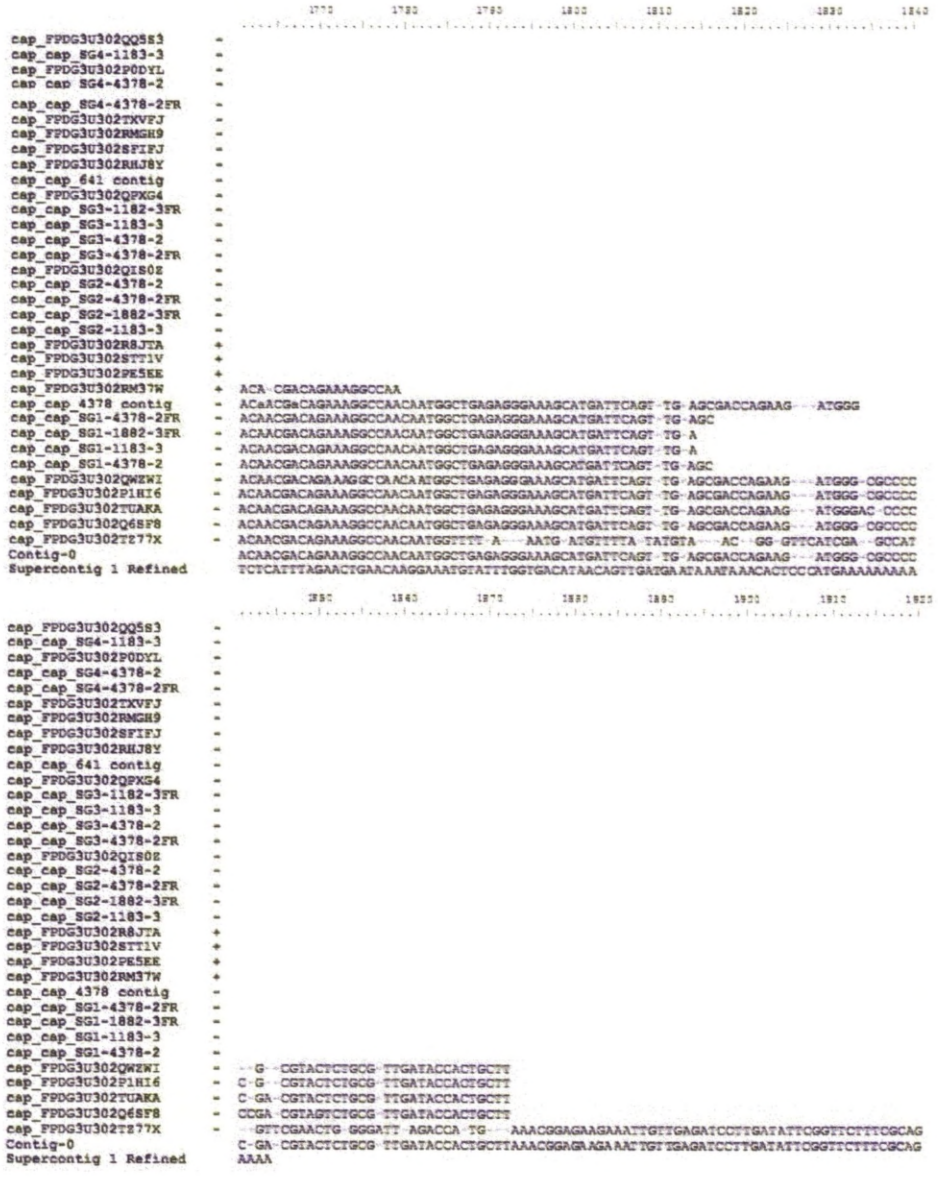

1290 1300 1310 1320 1330 1340 1350 1360
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR -
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z -
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR - GATGACAGTG CTCTCGTCATCAAAATCATAT CT AGAGCGA TGAGGATCATGGCGGT
cap_cap_SG2-1183-3 - GATGACAGTG CTCTCGTCATCAAAATCATAT CT AGAGCGA TGAGGATCATGGCGGTCGTAGACAAC
cap_FPDG3U302R8JTA + GATGACAGTG CTCTCGTCATCAAAATCATAT CT AGAGCGA TGAGGATCATGGCGGTCGTAGACAACCATGGCGCCT
cap_FPDG3U302STT1V + GATGACAGTGTCTCTCGTCATCAAAATCATAT CT AGAGCGAATGAGGATCATGGCGGTCGTAGACAACCATGGCGCCT
cap_FPDG3U302PE5EE + GATGACAGTG CTCTCGTCATCAAAATCATAT CT AGAGCGA TGAGGATCATGGCGGTCGTAGACAACCATGGCGCCT
cap_FPDG3U302RMJ7W + TG CTCTCGTCATCAAAATCATATTCTTAGAGCGA TGAGGATCATGGCGGTCGTAGACAACCATGGCGCCT
cap_cap_4378 contig - GATGACAGTG CTCTCGTCATCaAAATCATAt CT AGaGCGA TGAGGATCATGGCGGTCGTAGACAACCATGGCGCCT
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI -
cap_FPDG3U302P1RI6 -
cap_FPDG3U302TUAKA -
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302T277X -
Contig-0 GATGACAGTG CTCTCGTCATCAAAATCATAT CT AGAGCGA TGAGGATCATGGCGGTCGTAGACAACCATGGCGCCT
Supercontig 1 Refined ATCATGAGTGGTGCAATGACCGCGGCGTCTTTGCTTCCTGCCGGCTCGGTCTACGTCTTGCCACCTAATAAAGTTGTCGA

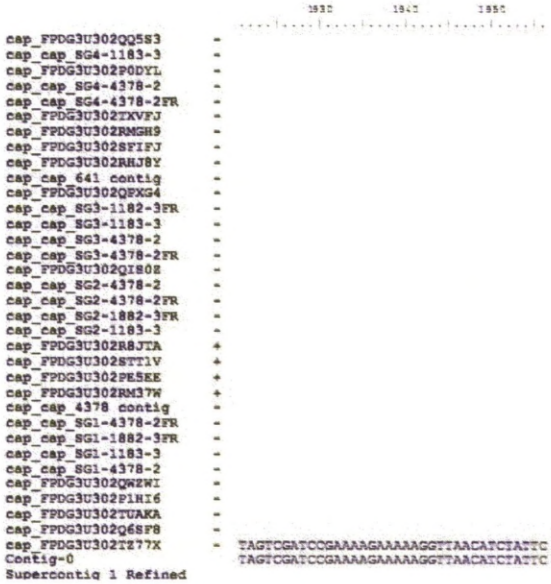

1370 1380 1390 1400 1410 1420 1430 1440
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR -
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z -
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR -
cap_cap_SG2-1183-3 -
cap_FPDG3U302R8JTA + CGAAGG CCATCGCAATACTGCGTTGAAAAATGGGA TGATACCAGAACGT
cap_FPDG3U302STT1V + CGAAGG CCATCGCAATACTGCGTTGAAAAATGGGAGTGATACCAGAACGTCCCCGCTTGATCTGGGACGCTAAACTTGT
cap_FPDG3U302PE5EE + CGAAGG CCATCGCAATACTGCGTTGAAAAATGGGAGTGATACCAAAACGTCCCCGCTTGATCTGGGACGCTAAACTTGT
cap_FPDG3U302RMJ7W + CGAAGG CCATCGCAATACTGCGTTGAAAAATGGGAGTGATACCAAAACGTCCCCGCTCGATCTGGGACGCTAAACTTGT
cap_cap_4378 contig - CGAAGG CCATCGCAATACTGCGTTGaAAAATGGGAgTGATACCagaaCGTcCCCGCtTGATCTgGGACGCTaAACtTGT
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI - ACGATCGCAATACTGCGTTGAAAA TGGGAGTGATACCAGAACGTCCC GCTTGATCTGG ACGCTAA CT GT
cap_FPDG3U302P1RI6 -
cap_FPDG3U302TUAKA -
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302T277X -
Contig-0 CGAAGG CCATCGCAATACTGCGTTGAAAAATGGGAGTGATACCAGAACGTCCCCGCTTGATCTGGGACGCTAAACTTGT
Supercontig 1 Refined GATCTCAATGCCCGGAGGGGCGGTGGGCAGTCCTCATCCTATTCATTTGCACGGCCATGCATTCTCGGTTGTGCGAAGTG


1450 1460 1470 1480 1490 1500 1510 1520
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR -
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z -
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR -
cap_cap_SG2-1183-3 -
cap_FPDG3U302R8JTA +
cap_FPDG3U302STT1V + ACAAGAAGGAGT GGCC GGG CGATAGGGCACTGGTTGACGCCCACTGGACCATCGGCCCATGTAGTCCCTTCTTGGAA
cap_FPDG3U302PE5EE + ACAAGAAGGAGT GGCCCGGGGCGATAGGGCACTGGTTGACGCCCACTGGACCATCGGCCCATGTAGTCCCTTCTTGGAA
cap_FPDG3U302RM37W + ACAAGAAGGAGT GGCCCGGGGCGATAGGGCACTGGTTGACGCCCACTGGACTATCGGCCCATGTAGTCCCTTCTTGGAA
cap_cap_4378 contig - ACAAGAAGGAGT GGcCCgGGGCGATaGGGCACTGGTTGACGCCCACTGGAcCATCGGcCCATGTAGTCCCTTCTTGGAA
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI - ACAAGAAGGAGTCGGCCCGGGGCGATAGGGCACTGGTTGACGCCCACTGGACCATCGGCC ATGTAGTCCCTTCTTGGAA
cap_FPDG3U302P1HI6 - GGAGT GGCCCGGG CGATAGGGCACTGGTTGACGCCCACTGGACCATCGGCC ATGTAGTCCCTTCTTGGAA
cap_FPDG3U302TUAKA - TTCTTGGAA
cap_FPDG3U302Q6SF8 -
cap_FPDG3U302TZ77X -
Contig-0 ACAAGAAGGAGT GGCCCGGGGCGATAGGGCACTGGTTGACGCCCACTGGACCATCGGCCCATGTAGTCCCTTCTTGGAA
Supercontig 1 Refined CAGGAAGCTCTGTGTATAATTACGCAAACCCCGTTCAGCGTGATGTCGTCAGTATTGGTGTGGCTGGCGATAACACGACG

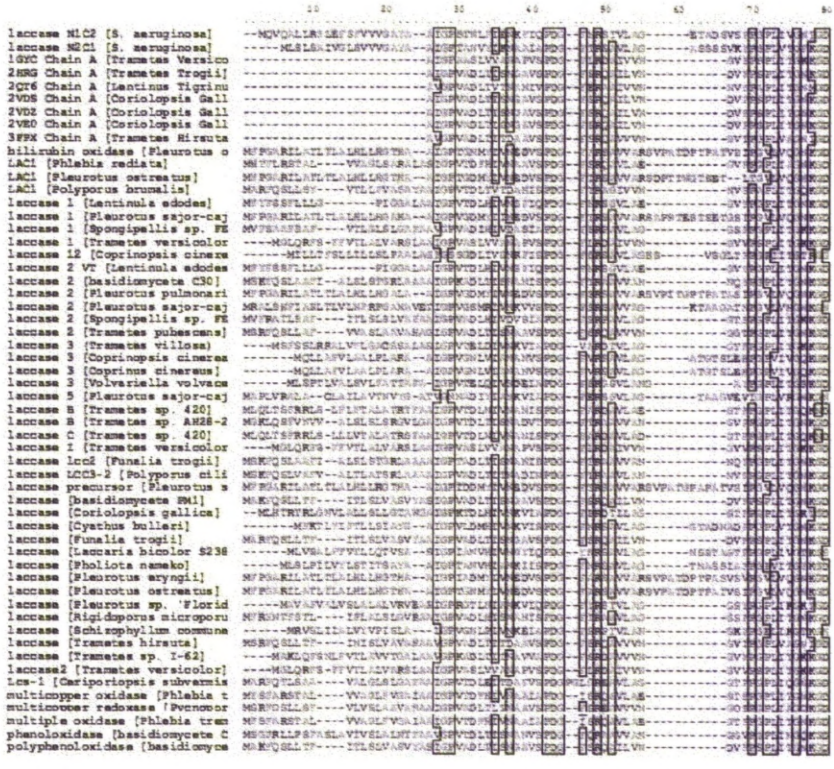

1530 1540 1550 1560 1570 1580 1590 1600
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR -
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z -
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR -
cap_cap_SG2-1183-3 -
cap_FPDG3U302R8JTA +
cap_FPDG3U302STT1V + CACA CC GTGCCAA TGAATACTTGTGCTCAGCAACATGGTGTTATCGGTGAGCTGGTCAATA CATTC
cap_FPDG3U302PE5EE + CATA CC GTGCCAA TGAATACTTGTGCTCAGCAACATGGTGTTATCGGTGAGCTGGTCAATAACATTCAAGCTAAAT
cap_FPDG3U302RM37W + CATA CC GTGCCAA TGAATACTTGTGCTCAGCAACATGGTGTTATCGGTGAGCTGGTCAATAACATTCAAGCTAAAT
cap_cap_4378 contig - CAtA CC GTGCCAA TGAATACTTGTGCTCAGCAACATGGTGTTATCGGTGAGCTGGTCAATaACATTCAAGCTAAAT
cap_cap_SG1-4378-2FR - AAGC
cap_cap_SG1-1882-3FR - C
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI - CACA CC GCGCCAA TGAATACTTGTGCTCAGCAACATGGTGTTATCGGTGAGCTGGTCAATAACATTCAAGCTAAAT
cap_FPDG3U302P1HI6 - CATAACCCGTGCCAAATGAATACTTGTGCTCAGCAACATGGTGTTATCGGTGAGCTGGTCAATAACATTCAAGCTAAAT
cap_FPDG3U302TUAKA - CATA CC GTGCCAA TGAATACTTGTGCTCAGCAACATGGTGTTATCGGTGAGCTGGTCAATAACATTCAAGCTAAAT
cap_FPDG3U302Q6SF8 - TGAATACTTGTGCTCAGCAACATGGTGT ATCGGTGAGCTGGTCAATAACATTCAAGCTAAAT
cap_FPDG3U302TZ77X -
Contig-0 CATA CC GTGCCAA TGAATACTTGTGCTCAGCAACATGGTGTTATCGGTGAGCTGGTCAATAACATTCAAGCTAAAT
Supercontig 1 Refined ATTCGCTTCAAGACAGATAATTCTGGTCCATGGATCATGCATTGCCATATCGACTGGCATCTTGAGATCGGCCTTGCAGT

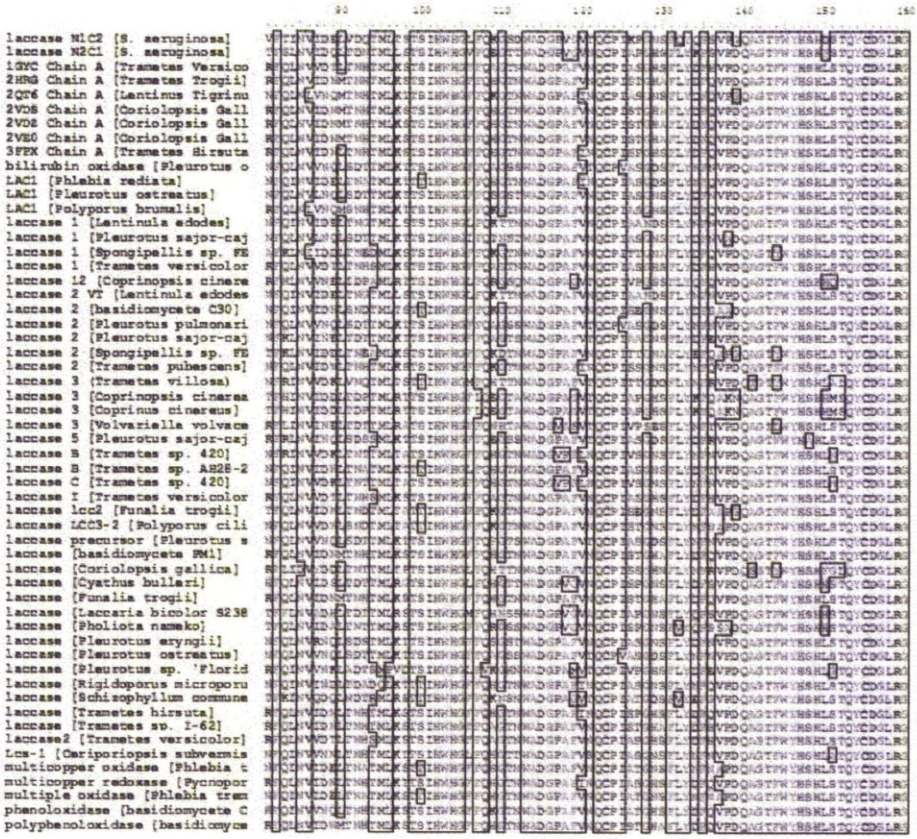

cap_FPDG3U302QQ5S3
cap_cap_SG4-1183-3
cap_FPDG3U302P0DYL
cap_cap_SG4-4378-2
cap_cap_SG4-4378-2FR
cap_FPDG3U302TXVFJ
cap_FPDG3U302RMGH9
cap_FPDG3U302SFIFJ
cap_FPDG3U302RHJ8Y
cap_cap_641_contig
cap_cap_SG3-4378-2
cap_cap_SG3-4378-2FR
cap_FPDG3U302QI80Z
cap_cap_SG2-4378-2
cap_cap_SG2-4378-2FR
cap_cap_SG2-1882-3FR
cap_cap_SG2-1183-3
cap_FPDG3U302R8JTA
cap_FPDG3U302STT1V
cap_FPDG3U302PE5EE
cap_FPDG3U302RM37W
cap_cap_4378_contig
cap_cap_SG1-4378-2FR
cap_cap_SG1-1882-3FR
cap_cap_SG1-1183-3
cap_cap_SG1-4378-2
cap_FPDG3U302QWZWI
cap_FPDG3U302P1HI6
cap_FPDG3U302TUAKA
cap_FPDG3U302Q6SF8
cap_FPDG3U302TZ77X
Contig-0
Supercontig_1_Refined

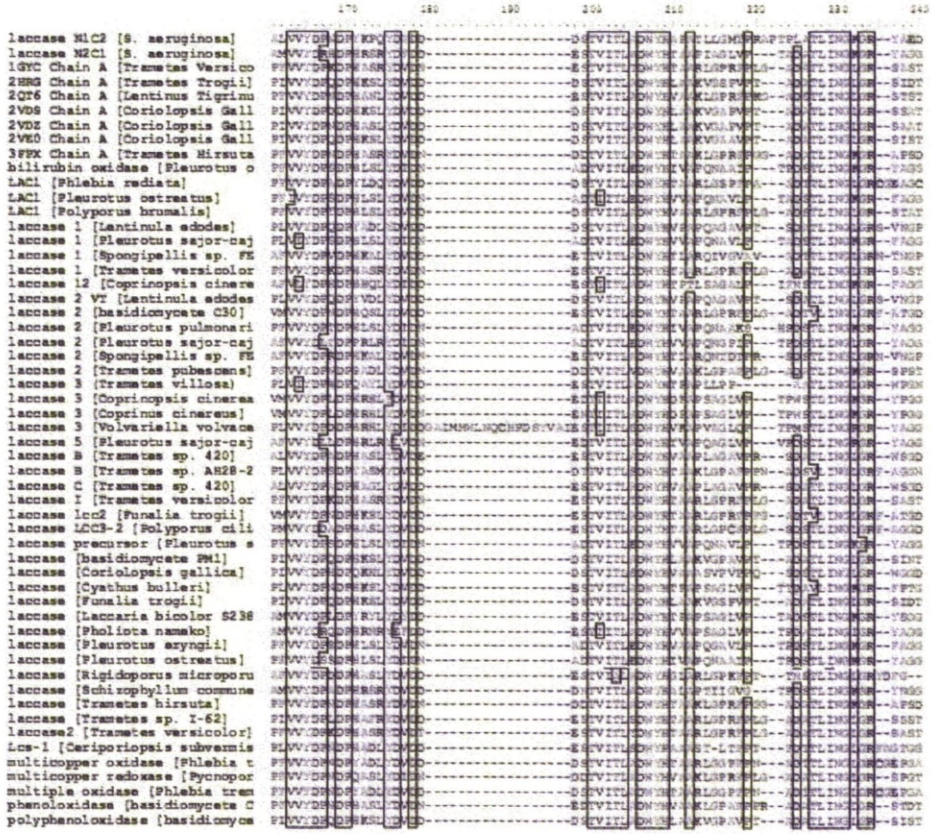

1770 1780 1790 1800 1810 1820 1830 1840
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR -
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z -
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR -
cap_cap_SG2-1183-3 -
cap_FPDG3U302R8JTA +
cap_FPDG3U302STT1V +
cap_FPDG3U302PE5EE +
cap_FPDG3U302RM37W + ACA CGACAGAAAGGCCAA
cap_cap_4378 contig - ACaACGaCAGAAAGGCCAACAATGGCTGAGAGGGAAAGCATGATTCAGT TG AGCGACCAGAAG ATGGG
cap_cap_SG1-4378-2FR - ACAACGACAGAAAGGCCAACAATGGCTGAGAGGGAAAGCATGATTCAGT TG AGC
cap_cap_SG1-1882-3FR - ACAACGACAGAAAGGCCAACAATGGCTGAGAGGGAAAGCATGATTCAGT TG A
cap_cap_SG1-1183-3 - ACAACGACAGAAAGGCCAACAATGGCTGAGAGGGAAAGCATGATTCAGT TG A
cap_cap_SG1-4378-2 - ACAACGACAGAAAGGCCAACAATGGCTGAGAGGGAAAGCATGATTCAGT TG AGC
cap_FPDG3U302QWZWI - ACAACGACAGAAAGCCAACAATGGCTGAGAGGGAAAGCATGATTCAGT TG AGCGACCAGAAG ATGGG CGCCCC
cap_FPDG3U302P1HI6 - ACAACGACAGAAAGGCCAACAATGGCTGAGAGGGAAAGCATGATTCAGT TG AGCGACCAGAAG ATGGG CGCCCC
cap_FPDG3U302TUAKA - ACAACGACAGAAAGGCCAACAATGGCTGAGAGGGAAAGCATGATTCAGT TG AGCGACCAGAAG ATGGGAC CCCC
cap_FPDG3U302Q6SF8 - ACAACGACAGAAAGGCCAACAATGGCTGAGAGGGAAAGCATGATTCAGT TG AGCGACCAGAAG ATGGG CGCCCC
cap_FPDG3U302TZ77X - ACAACGACAGAAAGGCCAACAATGGTTTT A AATG ATGTTTA TATGTA AC GG GTTCATCGA GCCAT
Contig-0 ACAACGACAGAAAGGCCAACAATGGCTGAGAGGGAAAGCATGATTCAGT TG AGCGACCAGAAG ATGGG CGCCCC
Supercontig 1 Refined TCTCATTTAGAACTGAACAAGGAAATGTATTTGGTGACATAACAGTTGATGAATAAATAAACACTCCCATGAAAAAAAAA
1860 1860 1870 1880 1890 1900 1910 1920
cap_FPDG3U302QQ5S3 -
cap_cap_SG4-1183-3 -
cap_FPDG3U302P0DYL -
cap_cap_SG4-4378-2 -
cap_cap_SG4-4378-2FR -
cap_FPDG3U302TXVFJ -
cap_FPDG3U302RMGH9 -
cap_FPDG3U302SFIFJ -
cap_FPDG3U302RHJ8Y -
cap_cap_641 contig -
cap_FPDG3U302QPXG4 -
cap_cap_SG3-1182-3FR -
cap_cap_SG3-1183-3 -
cap_cap_SG3-4378-2 -
cap_cap_SG3-4378-2FR -
cap_FPDG3U302QIS0Z -
cap_cap_SG2-4378-2 -
cap_cap_SG2-4378-2FR -
cap_cap_SG2-1882-3FR -
cap_cap_SG2-1183-3 -
cap_FPDG3U302R8JTA +
cap_FPDG3U302STT1V +
cap_FPDG3U302PE5EE +
cap_FPDG3U302RM37W +
cap_cap_4378 contig -
cap_cap_SG1-4378-2FR -
cap_cap_SG1-1882-3FR -
cap_cap_SG1-1183-3 -
cap_cap_SG1-4378-2 -
cap_FPDG3U302QWZWI - G CGTACTCTGCG TTGATACCACTGCTT
cap_FPDG3U302P1HI6 - C G CGTACTCTGCG TTGATACCACTGCTT
cap_FPDG3U302TUAKA - C GA CGTACTCTGCG TTGATACCACTGCTT
cap_FPDG3U302Q6SF8 - CCGA CGTAGTCTGCG TTGATACCACTGCTT
cap_FPDG3U302TZ77X - GTTCGAACTG GGGATT AGACCA TG AAACGGAGAAGAAATTGTTGAGATCCTTGATATTCGGTTCTTTCGCAG
Contig-0 C GA CGTACTCTGCG TTGATACCACTGCTTAAACGGAGAAGAAACTGTTGAGATCCTTGATATTCGGTTCTTTCGCAG
Supercontig 1 Refined AAAA

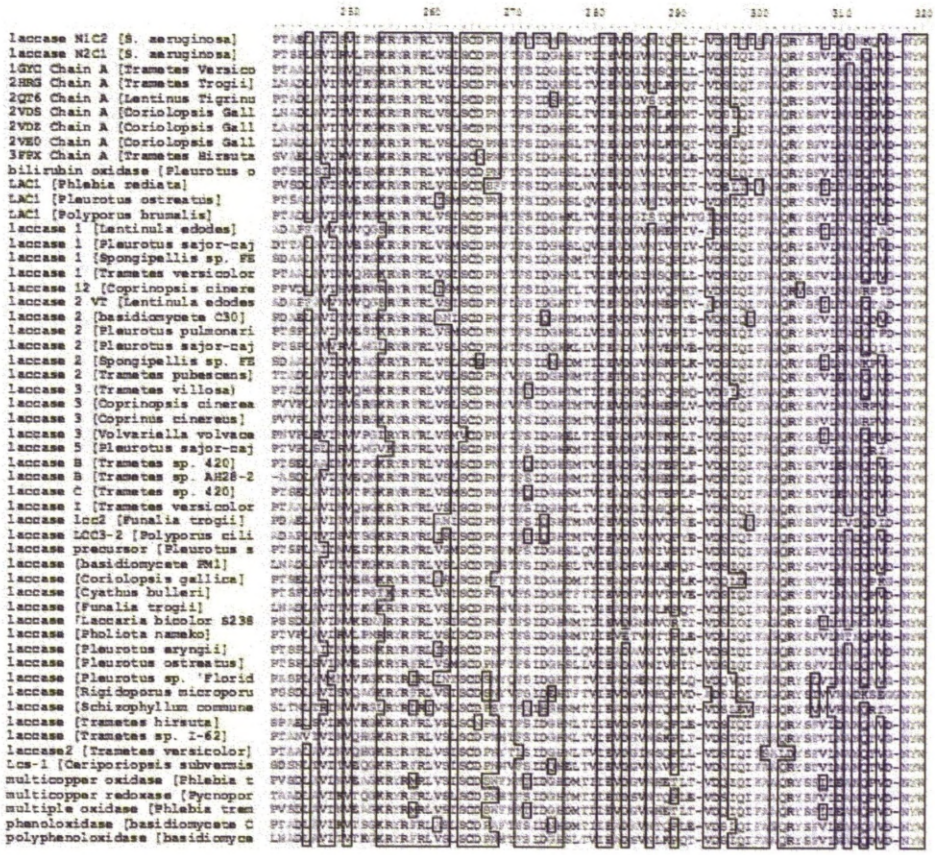


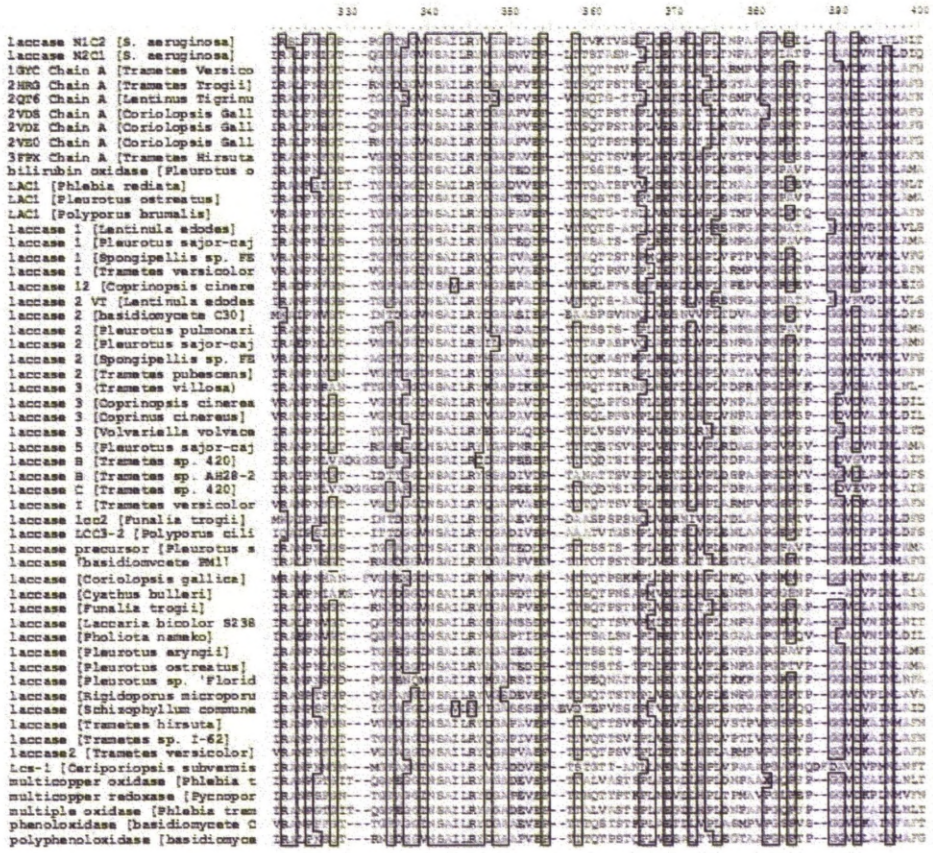
1.3. Multiple sequence alignment of translated protein sequences of *S. aeruginosa* laccases and their closest homologs.

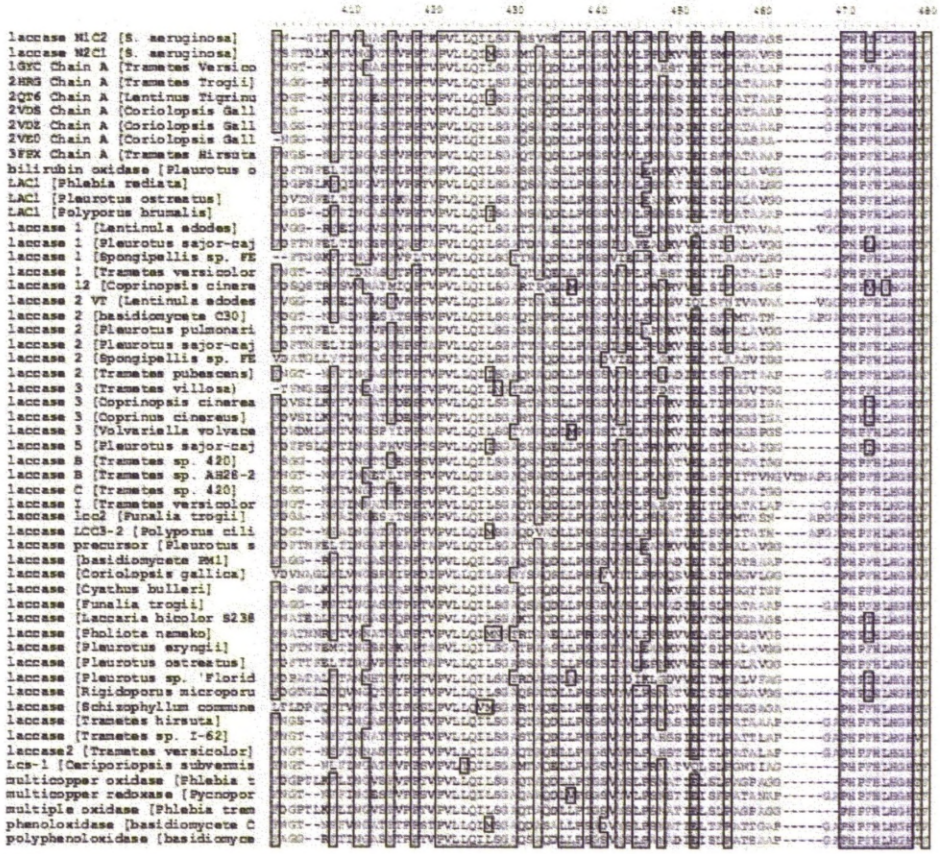


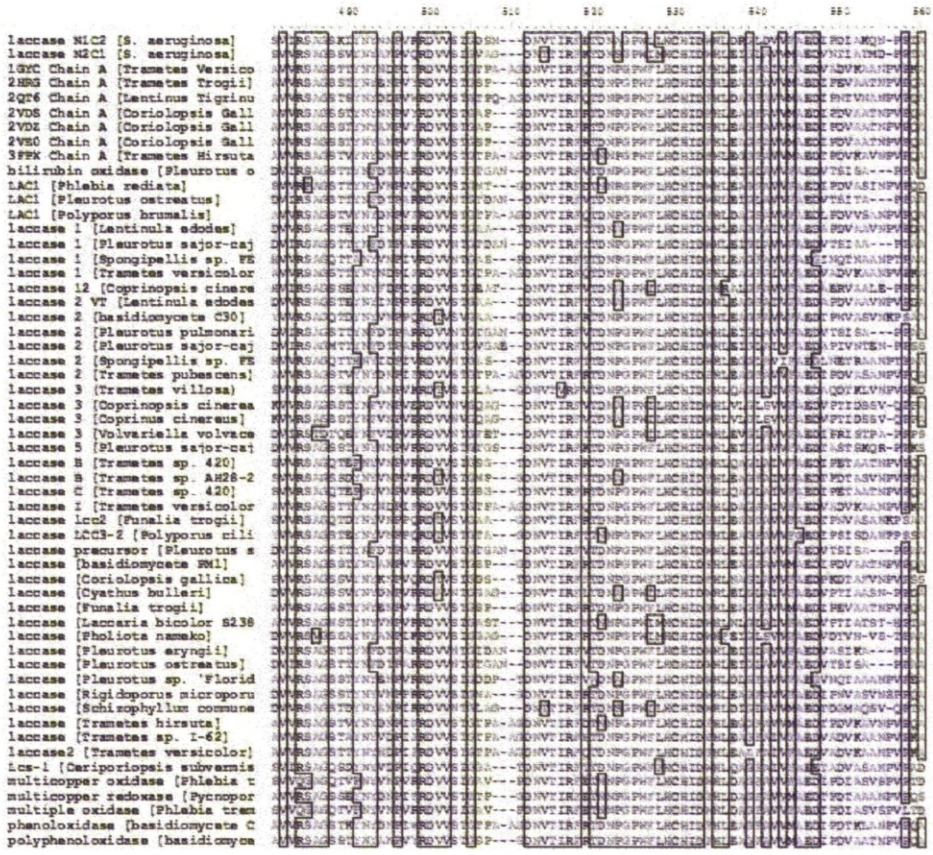

laccase N1C2 [S. aeruginosa]
laccase N2C1 [S. aeruginosa]
1GYC Chain A [Trametes Versico
2HRG Chain A [Trametes Trogii]
2QT6 Chain A [Lentinus Tigrinu
2VDS Chain A [Coriolopsis Gall
2VDZ Chain A [Coriolopsis Gall
2VE0 Chain A [Coriolopsis Gall
3FPX Chain A [Trametes Hirsuta
bilirubin oxidase [Pleurotus o
LAC1 [Phlebia radiata]
LAC1 [Pleurotus ostreatus]
LAC1 [Polyporus brumalis]
laccase 1 [Lentinula edodes]
laccase 1 [Pleurotus sajor-caj
laccase 1 [Spongipellis sp. FE
laccase 1 [Trametes versicolor
laccase L2 [Coprinopsis cinere
laccase 2 VT [Lentinula edodes
laccase 2 [basidiomycete C30]
laccase 2 [Pleurotus pulmonari
laccase 2 [Pleurotus sajor-caj
laccase 2 [Spongipellis sp. FE
laccase 2 [Trametes pubescens]
laccase 3 (Trametes villosa)
laccase 3 [Coprinopsis cinerea
laccase 3 [Coprinus cinereus]
laccase 3 [Volvariella volvace
laccase 5 [Pleurotus sajor-caj
laccase B [Trametes sp. 420]
laccase B [Trametes sp. AH28-2
laccase C [Trametes sp. 420]
laccase I [Trametes versicolor
laccase lcc2 [Funalia trogii]
laccase LCC3-2 [Polyporus cili
laccase precursor [Pleurotus s
laccase [basidiomycete PM1]
laccase [Coriolopsis gallica]
laccase [Cyathus bulleri]
laccase [Funalia trogii]
laccase [Laccaria bicolor S238
laccase [Pholiota nameko]
laccase [Pleurotus eryngii]
laccase [Pleurotus ostreatus]
laccase [Pleurotus sp. 'Florid
laccase [Rigidoporus microporu
laccase [Schizophyllum commune
laccase [Trametes hirsuta]
laccase [Trametes sp. I-62]
laccase2 [Trametes versicolor]
Lcs-1 [Cariporiopsis subvermis
multicopper oxidase [Phlebia t
multicopper redoxase [Pycnopor
multiple oxidase [Phlebia tren
phenoloxidase [basidiomycete C
polyphenoloxidase [basidiomyce


170
180
190
200
210
220
230
240
laccase N1C2 [S. aeruginosa]
laccase N2C1 [S. aeruginosa]
1GYC Chain A [Trametes Versico
2HRG Chain A [Trametes Trogii]
2QT6 Chain A [Lentinus Tigrinu
2VD8 Chain A [Coriolopsis Gall
2VDZ Chain A [Coriolopsis Gall
2VE0 Chain A [Coriolopsis Gall
3FPX Chain A [Trametes Hirsuta
bilirubin oxidase [Pleurotus o
LAC1 [Phlebia radiata]
LAC1 [Pleurotus ostreatus]
LAC1 [Polyporus brumalis]
laccase 1 [Lentinula edodes]
laccase 1 [Pleurotus sajor-caj
laccase 1 [Spongipellis sp. FE
laccase 1 [Trametes versicolor
laccase 12 [Coprinopsis cinere
laccase 2 VI [Lentinula edodes
laccase 2 [basidiomycete C30]
laccase 2 [Pleurotus pulmonari
laccase 2 [Pleurotus sajor-caj
laccase 2 [Spongipellis sp. FE
laccase 2 [Trametes pubescens]
laccase 3 [Trametes villosa]
laccase 3 [Coprinopsis cinerea
laccase 3 [Coprinus cinereus]
laccase 3 [Volvariella volvace
laccase 5 [Pleurotus sajor-caj
laccase B [Trametes sp. 420]
laccase B [Trametes sp. AH28-2
laccase C [Trametes sp. 420]
laccase I [Trametes versicolor
laccase lcc2 [Funalia trogii]
laccase LCC3-2 [Polyporus cili
laccase precursor [Pleurotus s
laccase [basidiomycete PM1]
laccase [Coriolopsis gallica]
laccase [Cyathus bulleri]
laccase [Funalia trogii]
laccase [Laccaria bicolor S238
laccase [Pholiota namako]
laccase [Pleurotus eryngii]
laccase [Pleurotus ostreatus]
laccase [Rigidoporus microporu
laccase [Schizophyllum commune
laccase [Trametes hirsuta]
laccase [Trametes sp. I-62]
laccase2 [Trametes versicolor]
Lcs-1 [Ceriporiopsis subvermis
multicopper oxidase [Phlebia t
multicopper radoxase [Pycnopor
multiple oxidase [Phlebia tren
phenoloxidase [basidiomycete C
polyphenoloxidase [basidiomyce

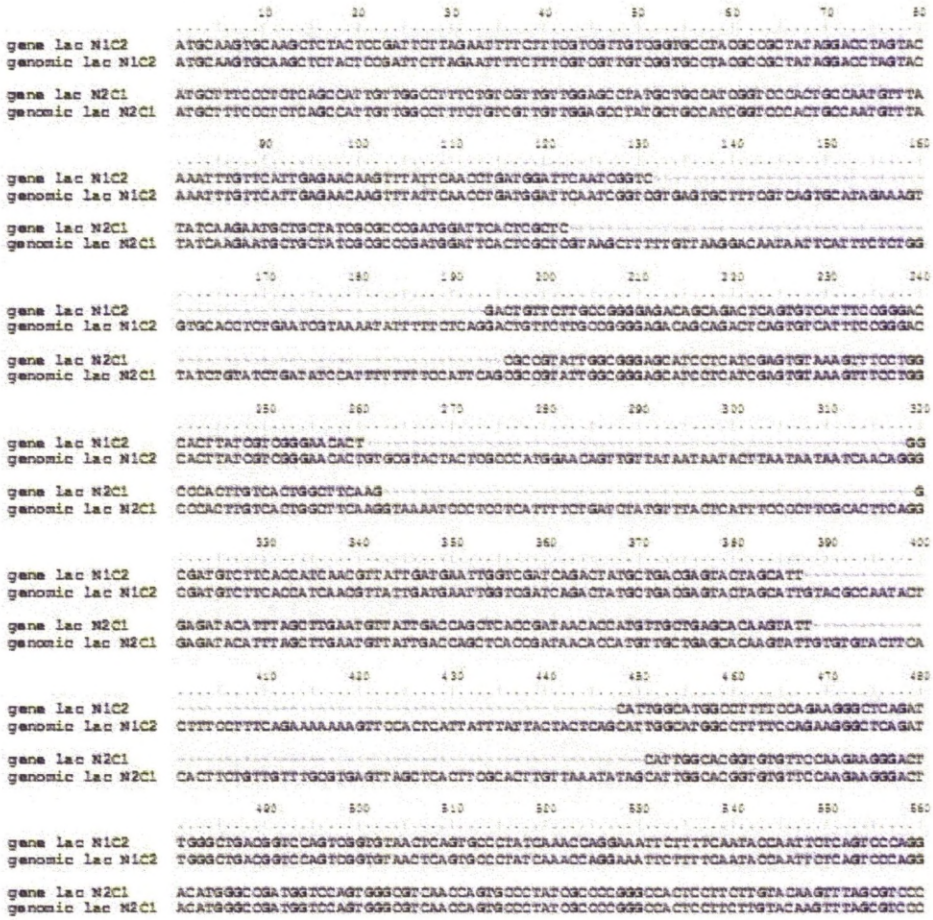

laccase N1C2 [S. aeruginosa]
laccase N2C1 [S. aeruginosa]
1GYC Chain A [Trametes Versico
2HRG Chain A [Trametes Trogii]
2QT6 Chain A [Lentinus Tigrinu
2VDS Chain A [Coriolopsis Gall
2VDZ Chain A [Coriolopsis Gall
2VE0 Chain A [Coriolopsis Gall
3FPX Chain A [Trametes Hirsuta
bilirubin oxidase [Pleurotus o
LAC1 [Phlebia radiata]
LAC1 [Pleurotus ostreatus]
LAC1 [Polyporus brumalis]
laccase 1 [Lentinula edodes]
laccase 1 [Pleurotus sajor-caj
laccase 1 [Spongipellis sp. FE
laccase 1 [Trametes versicolor
laccase 12 [Coprinopsis cinere
laccase 2 VT [Lentinula edodes
laccase 2 [basidiomycete C30]
laccase 2 [Pleurotus pulmonari
laccase 2 [Pleurotus sajor-caj
laccase 2 [Spongipellis sp. FE
laccase 2 [Trametes pubescens]
laccase 3 [Trametes villosa]
laccase 3 [Coprinopsis cinerea
laccase 3 [Coprinus cinereus]
laccase 3 [Volvariella volvace
laccase 5 [Pleurotus sajor-caj
laccase B [Trametes sp. 420]
laccase B [Trametes sp. AH28-2
laccase C [Trametes sp. 420]
laccase I [Trametes versicolor
laccase lcc2 [Funalia trogii]
laccase LCC3-2 [Polyporus cili
laccase precursor [Pleurotus s
laccase [basidiomycete RM1]
laccase [Coriolopsis gallica]
laccase [Cyathus bulleri]
laccase [Funalia trogii]
laccase [Laccaria bicolor S238
laccase [Pholiota nameko]
laccase [Pleurotus eryngii]
laccase [Pleurotus ostreatus]
laccase [Pleurotus sp. 'Florid
laccase [Rigidoporus microporu
laccase [Schizophyllum commune
laccase [Trametes hirsuta]
laccase [Trametes sp. I-62]
laccase2 [Trametes versicolor]
Lcs-1 [Ceriporiopsis subvermis
multicopper oxidase [Phlebia t
multicopper redoxase [Pycnopor
multiple oxidase [Phlebia tren
phenoloxidase [basidiomycete C
polyphenoloxidase [basidiomyce

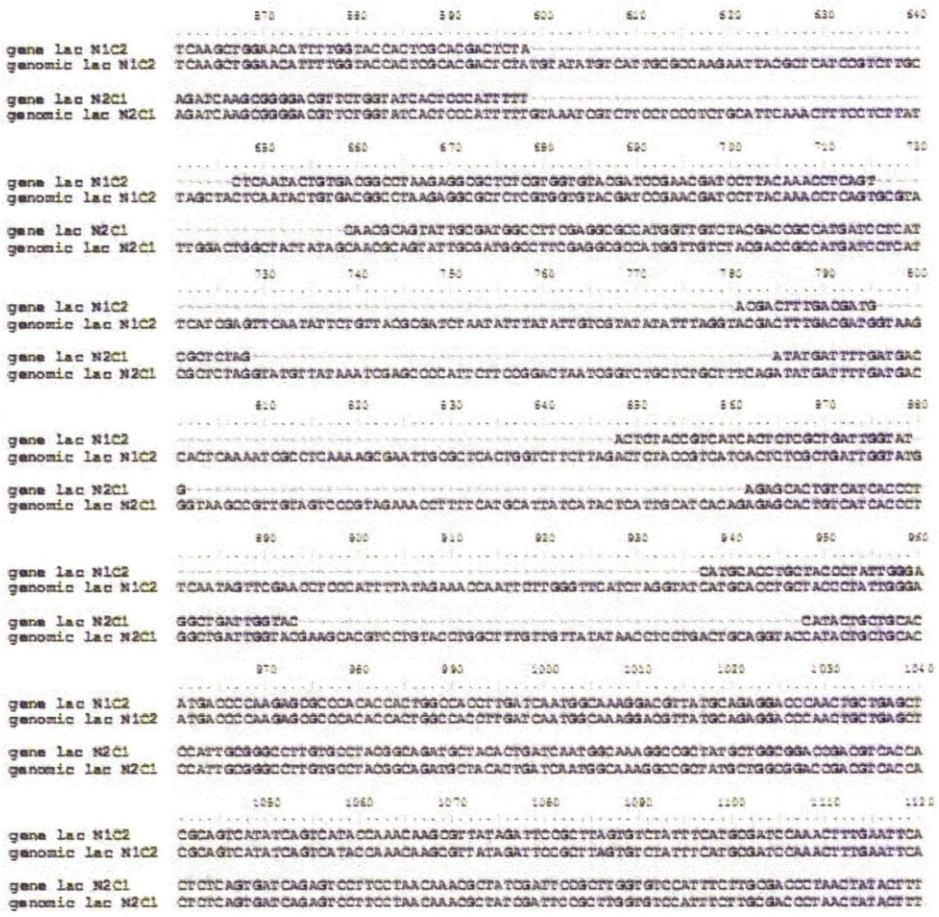

330
340
350
360
370
380
390
400
laccase N1C2 [S. aeruginosa]
laccase N2C1 [S. aeruginosa]
1GYC Chain A [Trametes Versico
2HRG Chain A [Trametes Trogii]
2QT6 Chain A [Lentinus Tigrinu
2VDS Chain A [Coriolopsis Gall
2VDZ Chain A [Coriolopsis Gall
2VZ0 Chain A [Coriolopsis Gall
3FPX Chain A [Trametes Hirsuta
bilirubin oxidase [Pleurotus o
LAC1 [Phlebia radiata]
LAC1 [Pleurotus ostreatus]
LAC1 [Polyporus brumalis]
laccase 1 [Lentinula edodes]
laccase 1 [Pleurotus sajor-caj
laccase 1 [Spongipellis sp. FE
laccase 1 [Trametes versicolor
laccase 12 [Coprinopsis cinere
laccase 2 VT [Lentinula edodes
laccase 2 [basidiomycete C30]
laccase 2 [Pleurotus pulmonari
laccase 2 [Pleurotus sajor-caj
laccase 2 [Spongipellis sp. FE
laccase 2 [Trametes pubescens]
laccase 3 (Trametes villosa)
laccase 3 [Coprinopsis cinerea
laccase 3 [Coprinus cinereus]
laccase 3 [Volvariella volvace
laccase 5 [Pleurotus sajor-caj
laccase B [Trametes sp. 420]
laccase B [Trametes sp. AH28-2
laccase C [Trametes sp. 420]
laccase I [Trametes versicolor
laccase lcc2 [Funalia trogii]
laccase LCC3-2 [Polyporus cili
laccase precursor [Pleurotus s
laccase [basidiomycete PM1]
laccase [Coriolopsis gallica]
laccase [Cyathus bulleri]
laccase [Funalia trogii]
laccase [Laccaria bicolor S238
laccase [Pholiota namako]
laccase [Pleurotus aryngii]
laccase [Pleurotus ostreatus]
laccase [Pleurotus sp. 'Florid
laccase [Rigidoporus microporu
laccase [Schizophyllum commune
laccase [Trametes hirsuta]
laccase [Trametes sp. I-62]
laccase2 [Trametes versicolor]
Lcs-1 [Ceriporiopsis subvermis
multicopper oxidase [Phlebia t
multicopper redoxase [Pycnopor
multiple oxidase [Phlebia trem
phenoloxidase [basidiomycete C
polyphenoloxidase [basidiomyce

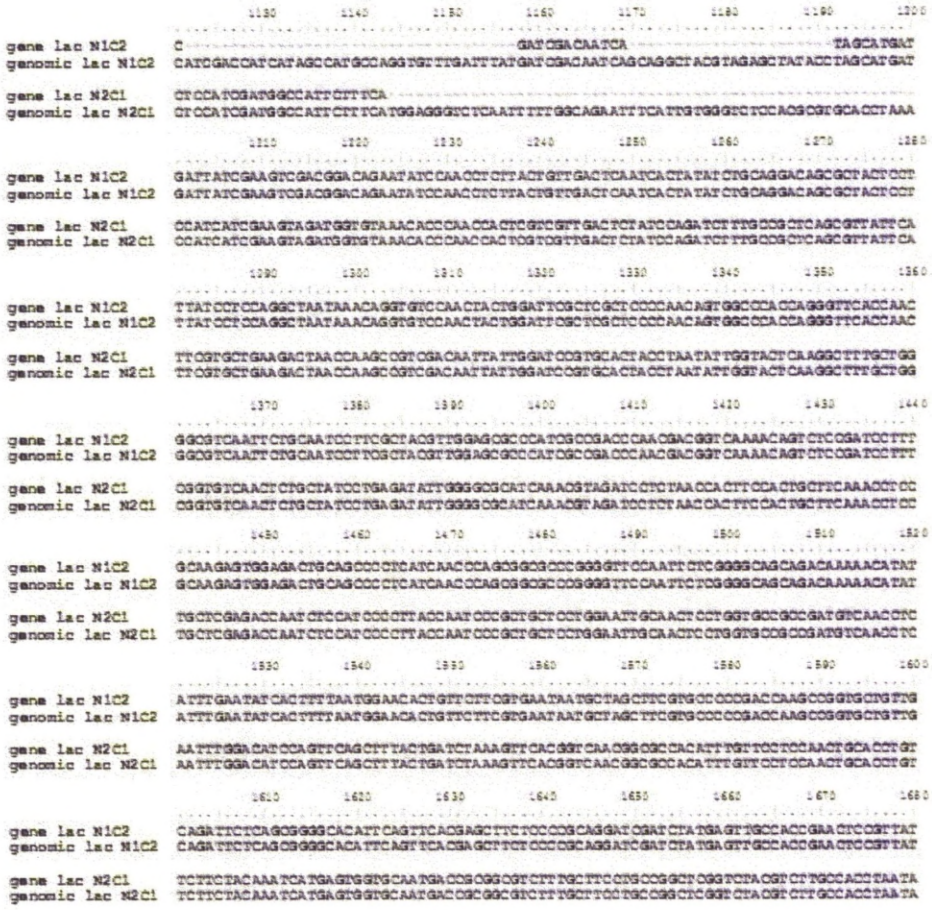

laccase N1C2 [S. aeruginosa]
laccase N2C1 [S. aeruginosa]
1GYC Chain A [Trametes Versico
2HRG Chain A [Trametes Trogii]
2QT6 Chain A [Lentinus Tigrinu
2VDS Chain A [Coriolopsis Gall
2VDZ Chain A [Coriolopsis Gall
2VE0 Chain A [Coriolopsis Gall
3FBX Chain A [Trametes Hirsuta
bilirubin oxidase [Pleurotus o
LAC1 [Phlebia radiata]
LAC1 [Pleurotus ostreatus]
LAC1 [Polyporus brumalis]
laccase 1 [Lentinula edodes]
laccase 1 [Pleurotus sajor-caj
laccase 1 [Spongipellis sp. PE
laccase 1 [Trametes versicolor
laccase 12 [Coprinopsis cinere
laccase 2 VT [Lentinula edodes
laccase 2 [basidiomycete C30]
laccase 2 [Pleurotus pulmonari
laccase 2 [Pleurotus sajor-caj
laccase 2 [Spongipellis sp. PE
laccase 2 [Trametes pubescens]
laccase 3 (Trametes villosa)
laccase 3 [Coprinopsis cinerea
laccase 3 [Coprinus cinereus]
laccase 3 [Volvariella volvace
laccase 5 [Pleurotus sajor-caj
laccase B [Trametes sp. 420]
laccase B [Trametes sp. AH28-2
laccase C [Trametes sp. 420]
laccase I [Trametes versicolor
laccase lcc2 [Funalia trogii]
laccase LCC3-2 [Polyporus cili
laccase precursor [Pleurotus s
laccase [basidiomycete PM1]
laccase [Coriolopsis gallica]
laccase [Cyathus bulleri]
laccase [Funalia trogii]
laccase [Laccaria bicolor S238
laccase [Pholiota nameko]
laccase [Pleurotus eryngii]
laccase [Pleurotus ostreatus]
laccase [Pleurotus sp. 'Florid
laccase [Rigidoporus microporu
laccase [Schizophyllum commune
laccase [Trametes hirsuta]
laccase [Trametes sp. I-62]
laccase2 [Trametes versicolor]
Lcs-1 [Cariporiopsis subvermis
multicopper oxidase [Phlebia t
multicopper redoxase [Pycnopor
multiple oxidase [Phlebia tram
phenoloxidase [basidiomycete C
polyphenoloxidase [basidiomyce

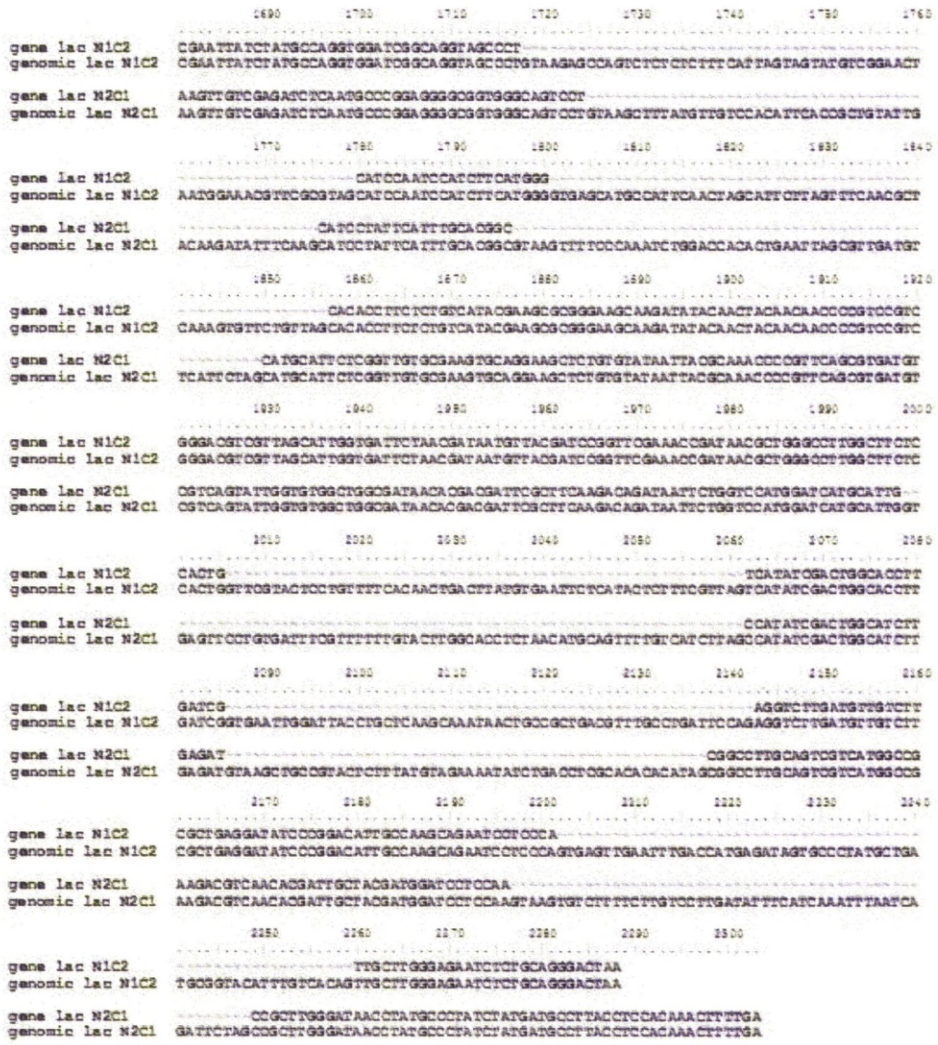

laccase N1C2 [S. aeruginosa]
laccase N2C1 [S. aeruginosa]
1GYC Chain A [Trametes Versico
2HRG Chain A [Trametes Trogii]
2QT6 Chain A [Lentinus Tigrinu
2VDS Chain A [Coriolopsis Gall
2VDZ Chain A [Coriolopsis Gall
2VS0 Chain A [Coriolopsis Gall
3FPX Chain A [Trametes Hirsuta
bilirubin oxidase [Pleurotus o
LAC1 [Phlebia radiata]
LAC1 [Pleurotus ostreatus]
LAC1 [Polyporus brumalis]
laccase 1 [Lentinula edodes]
laccase 1 [Pleurotus sajor-caj
laccase 1 [Spongipellis sp. FE
laccase 1 [Trametes versicolor
laccase 12 [Coprinopsis cinere
laccase 2 VT [Lentinula edodes
laccase 2 [basidiomycete C30]
laccase 2 [Pleurotus pulmonari
laccase 2 [Pleurotus sajor-caj
laccase 2 [Spongipellis sp. FE
laccase 2 [Trametes pubescens]
laccase 3 (Trametes villosa)
laccase 3 [Coprinopsis cinerea
laccase 3 [Coprinus cinereus]
laccase 3 [Volvariella volvace
laccase 5 [Pleurotus sajor-caj
laccase B [Trametes sp. 420]
laccase B [Trametes sp. AH28-2
laccase C [Trametes sp. 420]
laccase I [Trametes versicolor
laccase lcc2 [Funalia trogii]
laccase LCC3-2 [Polyporus cili
laccase precursor [Pleurotus s
laccase [basidiomycete RM1]
laccase [Coriolopsis gallica]
laccase [Cyathus bulleri]
laccase [Funalia trogii]
laccase [Laccaria bicolor S238
laccase [Pholiota nameko]
laccase [Pleurotus eryngii]
laccase [Pleurotus ostreatus]
laccase [Pleurotus sp. 'Florid
laccase [Rigidoporus microporu
laccase [Schizophyllum commune
laccase [Trametes hirsuta]
laccase [Trametes sp. I-62]
laccase2 [Trametes versicolor]
lcs-1 [Cariporiopsis subvermis
multicopper oxidase [Phlebia t
multicopper redoxase [Pycnopor
multiple oxidase [Phlebia tren
phenoloxidase [basidiomycete C
polyphenoloxidase [basidiomyce

1.4. Exons and introns of *S. aeruginosa* laccases N1C2 and N2C1.




gene lac N1C2
genomic lac N1C2
gene lac N2C1
genomic lac N2C1

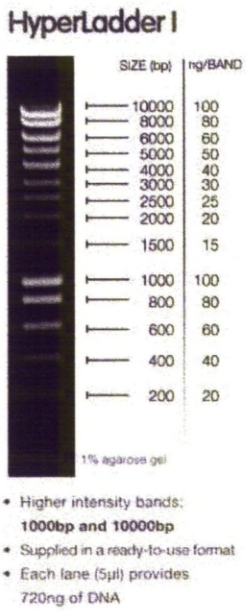

gene lac N1C2
genomic lac N1C2
gene lac N2C1
genomic lac N2C1


1690 1700 1710 1720 1730 1740 1750 1760
gene lac N1C2 CGAATTATCTATGCCAGGTGGATCGGCAGGTAGCCCT
genomic lac N1C2 CGAATTATCTATGCCAGGTGGATCGGCAGGTAGCCCTGTAAGAGCCAGTCTCTCTCTTTCATTAGTAGTATGTCGGAACT
gene lac N2C1 AAGTTGTCGAGATCTCAATGCCCGGAGGGGCGGTGGGCAGTCCT
genomic lac N2C1 AAGTTGTCGAGATCTCAATGCCCGGAGGGGCGGTGGGCAGTCCTGTAAGCTTTATGTTGTCCACATTCACCGCTGTATTG
1770 1780 1790 1800 1810 1820 1830 1840
gene lac N1C2 CATCCAATCCATCTTCATGGG
genomic lac N1C2 AATGGAAACGTTCGCGTAGCATCCAATCCATCTTCATGGGGTGAGCATGCCATTCAACTAGCATTCTTAGTTTCAACGCT
gene lac N2C1 CATCCTATTCATTTGCACGGC
genomic lac N2C1 ACAAGATATTTCAAGCATCCTATTCATTTGCACGGCGTAAGTTTTCCCAAATCTGGACCACACTGAATTAGCGTTGATGT
1850 1860 1870 1880 1890 1900 1910 1920
gene lac N1C2 CACACCTTCTCTGTCATACGAAGCGCGGGAAGCAAGATATACAACTACAACAACCCCGTCCGTC
genomic lac N1C2 CAAAGTGTTCTGTTAGCACACCTTCTCTGTCATACGAAGCGCGGGAAGCAAGATATACAACTACAACAACCCCGTCCGTC
gene lac N2C1 CATGCATTCTCGGTTGTGCGAAGTGCAGGAAGCTCTGTGTATAATTACGCAAACCCCGTTCAGCGTGATGT
genomic lac N2C1 TCATTCTAGCATGCATTCTCGGTTGTGCGAAGTGCAGGAAGCTCTGTGTATAATTACGCAAACCCCGTTCAGCGTGATGT
1930 1940 1950 1960 1970 1980 1990 2000
gene lac N1C2 GGGACGTCGTTAGCATTGGTGATTCTAACGATAATGTTACGATCCGGTTCGAAACCGATAACGCTGGGCCTTGGCTTCTC
genomic lac N1C2 GGGACGTCGTTAGCATTGGTGATTCTAACGATAATGTTACGATCCGGTTCGAAACCGATAACGCTGGGCCTTGGCTTCTC
gene lac N2C1 CGTCAGTATTGGTGTGGCTGGCGATAACACGACGATTCGCTTCAAGACAGATAATTCTGGTCCATGGATCATGCATTG
genomic lac N2C1 CGTCAGTATTGGTGTGGCTGGCGATAACACGACGATTCGCTTCAAGACAGATAATTCTGGTCCATGGATCATGCATTGGT
2010 2020 2030 2040 2050 2060 2070 2080
gene lac N1C2 CACTG TCATATCGACTGGCACCTT
genomic lac N1C2 CACTGGTTCGTACTCCTGTTTTCACAACTGACTTATGTGAATTCTCATACTCTTTCGTTAGTCATATCGACTGGCACCTT
gene lac N2C1 CCATATCGACTGGCATCTT
genomic lac N2C1 GAGTTCCTGTGATTTCGTTTTTTGTACTTGGCACCTCTAACATGCAGTTTTGTCATCTTAGCCATATCGACTGGCATCTT
2090 2100 2110 2120 2130 2140 2150 2160
gene lac N1C2 GATCG AGGTCTTGATGTTGTCTT
genomic lac N1C2 GATCGGTGAATTGGATTACCTGCTCAAGCAAATAACTGCCGCTGACGTTTGCCTGATTCCAGAGGTCTTGATGTTGTCTT
gene lac N2C1 GAGAT CGGCCTTGCAGTCGTCATGGCCG
genomic lac N2C1 GAGATGTAAGCTGCCGTACTCTTTATGTAGAAAATATCTGACCTCGCACACATAGCGGCCTTGCAGTCGTCATGGCCG
2170 2180 2190 2200 2210 2220 2230 2240
gene lac N1C2 CGCTGAGGATATCCCGGACATTGCCAAGCAGAATCCTCCCA
genomic lac N1C2 CGCTGAGGATATCCCGGACATTGCCAAGCAGAATCCTCCCAGTGAGTTGAATTGACCATGAGATAGTGCCCTATGCTGA
gene lac N2C1 AAGACGTCAACACGATTGCTACGATGGATCCTCCAA
genomic lac N2C1 AAGACGTCAACACGATTGCTACGATGGATCCTCCAAGTAAGTGTCTTTCTTGTCCTTGATATTTCATCAAATTTAATCA
2250 2260 2270 2280 2290 2300
gene lac N1C2 TTGCTTGGGAGAATCTCTGCAGGGACTAA
genomic lac N1C2 TGCGGTACATTTGTCACAGTTGCTTGGGAGAATCTCTGCAGGGACTAA
gene lac N2C1 CCGCTTGGGATAACCTATGCCCTATCTATGATGCCTTACCTCCACAAACTTTTGA
genomic lac N2C1 GATTCTAGCCGCTTGGGATAACCTATGCCCTATCTATGATGCCTTACCTCCACAAACTTTTGA

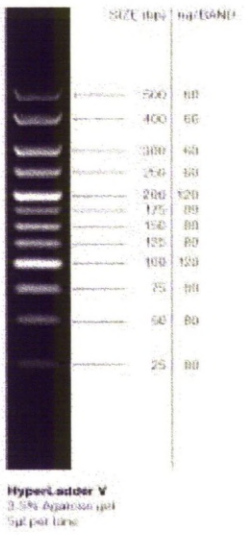
2. Appendix 2 – molecular weigth markers.

2.1. DNA markers.

2.1.1. Hyperladder I [Bioline].



2.1.2. Hyperladder V [Bioline].



2.2. Protein markers.

2.2.1. Prestained Protein Marker Plus [Fermentas].